Granger Centre for time series econometrics

Granger Centre Discussion Paper Series

Testing for a unit root in the presence of a possible break in trend

by

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

Granger Centre Discussion Paper No. 07/04

The University of Nottingham

Testing for a Unit Root in the Presence of a Possible Break in Trend[*]

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

[a]Department of Economics, University of Melbourne

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

University of Nottingham

October 2007

Abstract

We consider the issue of testing a time series for a unit root in the possible presence of a break in a linear deterministic trend at an unknown point in the series. We propose a new break fraction estimator which, where 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 our estimator converges to zero at rate $O_p(T^{-1/2})$. Used in conjunction with a quasi difference (QD) detrended unit root test that incorporates a trend break regressor, we show that these rates of convergence ensure that known break fraction null critical values are asymptotically valid. Unlike available procedures in the literature this holds even if there is no break in trend (the break fraction is zero). Here the trend break regressor is dropped from the deterministic component and standard QD detrended unit root test critical values then apply. We also propose a second procedure which makes use of a formal pre-test for a trend break in the series, including a trend break regressor only where the pre-test rejects the null of no break. Both procedures ensure that the correctly sized (near-) efficient unit root test that allows (does not allow) for a break in trend is applied in the limit when a trend break does (does not) occur.

Keywords: Unit root test; quasi difference de-trending; trend break; pre-test; asymptotic power.

JEL Classification: C22.

---

[*]Correspondence to: Robert Taylor, School of Economics, University of Nottingham, Nottingham NG7 2RD, UK. E-mail: `Robert.Taylor@nottingham.ac.uk`

1 Introduction

Testing for the presence of an autoregressive unit root process around a deterministic linear trend function has been an issue at the core of econometric research for the last quarter century. Since for many macroeconomic time series the possibility of a change in the underlying linear trend function at some point in the sample data needs to be entertained - for example, as might occur following a period of major economic upheaval or a political regime change - in the wake of the seminal paper by Perron (1989) it has become a matter of regular practice to apply a unit root test that allows for this kind of deterministic structural change in the trend function.

Perron (1989) treats the location of the potential trend break as known, *a priori*, to the practitioner. However, this assumption has attracted significant criticism (see, for example, Christiano, 1992) and, as a consequence, most recent approaches to this problem focus on the case where the possible break occurs at an unknown point in the sample which must be estimated in some way; see, *inter alia*, Zivot and Andrews (1992), Banerjee *et al.* (1992), Perron (1997) and Perron and Rodríguez (2003). This approach raises two obvious questions: first, how well can we estimate the break point when a break actually occurs and, second, how does the break point estimator behave when no break occurs? Both of these issues clearly have important forward implications for the behaviour of unit root tests that are based on estimated break points.

Taking the presence of a fixed trend in the data generation process (DGP) as a given, among augmented Dickey-Fuller (ADF) style unit root tests it is the Elliott *et al.* (1996) test based on quasi difference (QD) detrending that is near asymptotically efficient[1] when no additional broken trend is present. When a broken trend is also *known* to be present, Elliott *et al.*'s test is inconsistent and it is now a test based on Perron and Rodríguez's (2003) QD detrended ADF statistic which allows for a break in trend that is efficient. In this case, where the break occurs at a known point in the sample, the efficient test is analogous to Perron's (1989) test but using QD detrending. If the break occurs at an unknown point in the sample, the statistic needs to be evaluated based on an estimated break date, which must be consistent for the true break date at a sufficiently fast rate such that the critical values for the known break point case from Perron and Rodríguez (2003) are appropriate in the limit.[2] An alternative approach, considered for the case of OLS detrending by Zivot and Andrews (1992), and extended to the case of local GLS detrending by Perron and Rodríguez (2003), bases inference on the minimum of the ADF statistics calculated for all possible break dates within a given range.

Where a trend break does not occur the tests proposed in Perron and Rodríguez

---

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

[2]Perron and Rodríguez (2003,p.6) suggest one possible estimator based on the location of the maximum of a sequence of QD *t*-statistics for the presence of a trend break at each possible point within a trimmed set of points in the sample. Other authors have suggested the corresponding OLS estimator but, interestingly, this is not consistent at a sufficiently fast rate.

(2003) are not efficient, and indeed the efficiency losses can be quite substantial in such cases, as we demonstrate in this paper. It is obvious that this will be the case for the test which assumes a known possible break date, since a redundant trend break regressor will always be included. For the unknown break date case this also occurs because, in the absence of a trend break, the break point estimator they propose (in common with other currently available estimators) has a non-degenerate limit distribution over the range of possible break points from which it is calculated and, as such, will spuriously indicate the presence of a trend break.[3]

In practice, where it will be *unknown* as to whether a trend break occurs or not, this differing behaviour of the break point estimator also renders the *true* asymptotic critical values of the tests dependent on whether a break occurs or not. For the existing tests in the literature to be feasible in practice we are therefore faced with a choice: either, as in Perron and Rodríguez (2003), 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 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).

The aim of this paper is to rectify these drawbacks with the existing tests in the literature. We do so by proposing two approaches to testing for unit roots which allow for the possibility of a break in trend. Our first approach uses a new break point estimator which is a data-dependent modification of the estimator of the break point 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 modification drops out in the limit and the estimator collapses to the first difference estimator, which is shown to converge to the true point sufficiently rapidly such that the break point estimation error is negligible enough to allow the Perron and Rodríguez (2003) null critical values that are appropriate for a known break point to be applied in the limit. Secondly, it has the property that, when no break occurs, the estimator does not spuriously indicate a break point. In this case, in the limit, the modification forces our new estimator to put the estimated break point outside any range of break points considered to be feasible, so that it is the Elliott *et al.* (1996) statistic that would be applied, not the Perron and Rodríguez (2003) variant (which now incorporates a redundant trend break regressor). Again, this occurs sufficiently rapidly such that the Elliott *et al.* (1996) null critical values are relevant.

Our second approach builds upon earlier work by Kim and Perron (2006) in the context of OLS de-trended unit root tests which allow for a possible trend break at

---

[3]Similar problems arise with the minimum ADF-type tests but to a worse degree in that the location of the minimum of the ADF statistics is not even a consistent estimator of the true break date when a break occurs, and 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.

an unknown point in the sample, and by Harvey *et al.* (2007a) and Perron and Yabu (2007) for unit root tests where it is uncertain as to whether the data contain a constant and linear trend or simply a constant. What these approaches have in common is that they employ a pre-test for the presence of the deterministic component in question which is (asymptotically) robust as to whether the unit root holds or not, including the deterministic component in the subsequent unit root test only where the pre-test rejects. In the context of the present problem we show that this method achieves the same outcome in the limit as the first approach. As our trend break pre-test we use the recently developed test of Harvey *et al.* (2007b).[4] Where the trend break pre-test rejects, the Perron and Rodríguez (2003) variant of the unit root statistic is consequently employed, including a trend break dummy at the point in the sample identified by the (unmodified) first difference estimator, and where it does not the Elliott *et al.* (1996) variant is used. In order to ensure that the pre-test has no impact on the size of the resultant unit root test, the size of the pre-test is shrunk towards zero with the sample size, at a suitable rate. Our two proposed approaches are asymptotically equivalent and both ensure that an asymptotically correctly sized and (asymptotically) efficient unit root test which allows (does not allow) for a break in trend is applied in the limit when a trend break does (does not) occur.

After outlining our reference trend break model in section 2, in section 3 we develop a new break point (hereafter break fraction) estimator possessing the properties discussed above. Here, for unit root and near-unit root errors, we establish the rates of consistency of the new estimator for the true break fraction when a trend break occurs and show its limit behaviour under both errors when no break is present. At the same time, we show how the new estimator can be used in conjunction with the QD detrended unit root statistics described above and we establish the large sample behaviour of such a procedure. Corresponding results for our proposed pre-test-based approach are outlined in section 4. In section 5, we compare the asymptotic properties of our two proposed procedures with those of existing strategies based on the Zivot and Andrews (1992) and Perron and Rodríguez (2003) unit root tests. Here, we also conduct some comparison finite sample size and power simulations which, for the greater part, yield the same qualitative pattern as our asymptotic results. Concluding remarks are offered in section 6. Proofs of our results are contained in an Appendix.

In what follows we use the following notation: $\lfloor \cdot \rfloor$ to denote the integer part of its argument; '$\overset{p}{\to}$' and '$\overset{d}{\to}$' denote convergence in probability and weak convergence, respectively, in each case as the sample size diverges to positive infinity; $1(.)$ to denote the indicator function, and '$x := y$' indicates that x is defined by y.

---

[4]In recent independent research, a similar approach is used by Carrion-i-Silvestre *et al.* (2007) in the context of a model with multiple possible trend breaks, but using the trend break pre-test of Perron and Yabu (2006).

2   The Trend Break Model

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

$$y_t = \alpha_0 + \beta_0 t + \gamma_0 DT_t(\tau_0) + u_t, \quad t = 1, ..., T, \tag{1}$$

$$u_t = \rho_T u_{t-1} + \varepsilon_t, \quad t = 2, ..., T, \tag{2}$$

where $DT_t(\tau_0) := 1(t > \lfloor \tau_0 T \rfloor)(t - \lfloor \tau_0 T \rfloor)$ with $\lfloor \tau_0 T \rfloor$ the potential trend break point with associated break fraction τ_0. 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 no break is deemed allowable to occur. 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. It would also be possible to 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)-(2). However, as argued by Perron and Rodríguez (2003, pp.2,4), we need not analyze this case separately because a change in intercept is just a special case of a slowly evolving deterministic component (see Condition B of Elliott *et al.*, 1996, p.816) and, consequently, does not alter any of the large sample results presented in this paper.

The initial condition of the process is assumed to be such that $T^{-1/2} u_1 \xrightarrow{p} 0$, while the error process $\{\varepsilon_t\}$ in (2) is taken to satisfy the following conventional linear process assumption.

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

$$\varepsilon_t = c(L)\eta_t, \ c(L) := \sum_{j=0}^{\infty} c_j L^j$$

with $c(1)^2 > 0$ and $\sum_{i=0}^{\infty} i|c_i| < \infty$, and where $\{\eta_t\}$ is an IID sequence with mean zero, unit variance and finite fourth moment. The long run variance of ε_t is defined as $\omega_\varepsilon^2 := \lim_{T \to \infty} T^{-1} E(\sum_{t=1}^{T} \varepsilon_t)^2 = c(1)^2$.

Within (2), we set $\rho_T := 1 - c/T$ for $0 \leq c < \infty$ and we will be concerned with testing the unit root null hypothesis, $H_0 : c = 0$, against the local alternative, $H_1 : c > 0$.

3   Break Fraction Estimation and Unit Root Tests

In this section we discuss how asymptotically efficient unit root tests can be constructed in the presence of a (possible) break in trend. In section 3.1 we first consider the case where it is *known* that a break in trend has occurred and show that a QD de-trended test of the form considered in Perron and Rodríguez (2003) based around a first difference estimator of the (unknown) break fraction is efficient. In section 3.2 we then show that where, as will be the case in practice, it is unknown as to whether a trend break has occurred or not, that the approach outlined in section 3.1 no longer

delivers an efficient test. Here we suggest an approach based on a modification of the first difference estimator of the break fraction which is shown to deliver an efficient test both where a trend break occurs and where one does not.

3.1   The Case where a Break is Known to have Occurred

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 $\gamma_0 \neq 0$), we require an estimator of the unknown break fraction whose rate of consistency is sufficiently rapid for a unit root test based on that estimator to have an asymptotic null distribution that is the same as if the break fraction τ_0 were known. This requires that the estimator obtains a rate of consistency which is faster than $O_p(T^{-1/2})$.

As shown by Perron and Zhu (2005), the OLS estimator of τ_0,

$$\breve{\tau} := \arg\min_{\tau \in \Lambda} T^{-1} \sum_{t=1}^{T} \hat{u}_t(\tau)^2$$

where $\hat{u}_t(\tau)$, $t = 1, ..., T$ are the OLS residuals from a regression of y_t on $(1,\ t,\ DT_t(\tau))'$, does not have this property because it is is only $O_p\left(T^{-1/2}\right)$ consistent under $H_0 : c = 0$. However, as will be established in Lemma 1 below, by taking first differences of (1) an estimator with the required rate of consistency can be obtained. Specifically, we define our proposed first difference estimator of τ_0 as:

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

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

$$\Delta y_t = \beta_0 + \gamma_0 DU_t(\tau) + v_t, \tag{3}$$

where $DU_t(\tau) := 1(t > \lfloor \tau T \rfloor)$.

The regression model (3) represents a model for a mean shift in Δy_t, and the asymptotic properties of $\tilde{\tau}$ where v_t is a stationary and invertible linear process have been proved in Bai (1994). In particular, he showed that the break fraction estimator $\tilde{\tau}$ is $O_p\left(T^{-1}\right)$ consistent. His result applies in our case when $c = 0$ and so is relevant for our unit root test null distribution theory. The following lemma verifies that this rate also continues to hold when u_t contains a near-unit root (i.e. $c > 0$), which is important subsequently for establishing local alternative power functions of unit root tests based on this estimator.

Lemma 1 *Let y_t be generated according to (1) and (2) with $\rho_T = 1 - c/T$, $0 \leq c < \infty$, and let Assumption 1 hold. Then for the case where $\gamma_0 \neq 0$, $\tilde{\tau} = \tau_0 + O_p(T^{-1})$.*

Remark 1: It can be shown that the result stated in Lemma 1 also holds in the stable autoregressive case where $\rho_T = \rho$ with $|\rho| < 1$.

Remark 2: The rate of convergence stated in Lemma 1 can also be shown to hold for the corresponding QD estimator of τ_0 suggested in Perron and Rodríguez (2003, p.6), which we denote by $\hat{\tau}$ in what follows; cf. footnote 2.

Next consider the QD detrended ADF-type unit root test applied to (1) and (2). For known τ_0, the regression model (1) can be written as

$$y_t = X_t(\tau_0)' \theta_0 + u_t, \quad t = 1, ..., T, \tag{4}$$

where $X_t(\tau) = (1, t, DT_t(\tau))'$ and $\theta_0 = (\alpha_0, \beta_0, \gamma_0)'$. Applying a QD transformation to (4) yields

$$y_{\bar{c},t} = X_{\bar{c},t}(\tau_0)' \theta_0 + u_{\bar{c},t}, \tag{5}$$

where

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

$$X_{\bar{c},t}(\tau_0) := \begin{cases} X_1(\tau_0) & t = 1 \\ X_t(\tau_0) - \bar{\rho}_T X_{t-1}(\tau_0) & t = 2, \ldots, T \end{cases}$$

and $\bar{\rho}_T := 1 - \bar{c}/T$, where $\bar{c}$ is the QD parameter, which is generally chosen to be the value of c at which the asymptotic Gaussian local power envelope for a given significance level has power equal to 0.50. We define $\hat{\theta}_{\bar{c}}$ to be the OLS estimator in (5) and the residuals from (4) are then $\hat{u}_t := y_t - X_t(\tau_0)' \hat{\theta}_{\bar{c}}$. The QD detrended ADF test rejects for large negative values of the regression t-statistic for $\phi = 0$ in the ADF-type regression

$$\Delta \hat{u}_t = \phi \hat{u}_{t-1} + \sum_{j=1}^{p} \delta_j \Delta \hat{u}_{t-j} + e_{p,t}, \quad t = p + 2, ..., T. \tag{6}$$

We denote this statistic $ADF\text{-}GLS^{tb}(\tau_0, \bar{c})$.

For unknown τ_0, we simply repeat the preceding procedure but with τ_0 replaced by $\tilde{\tau}$ throughout. That is, we obtain the residuals $\tilde{u}_t := y_t - X_t(\tilde{\tau})' \tilde{\theta}_{\bar{c}}$, where $\tilde{\theta}_{\bar{c}}$ is the OLS estimator from a regression of $y_{\bar{c},t}$ on $X_{\bar{c},t}(\tilde{\tau})$, and then estimate the ADF-type regression

$$\Delta \tilde{u}_t = \phi \tilde{u}_{t-1} + \sum_{j=1}^{p} \delta_j \Delta \tilde{u}_{t-j} + e_{p,t}, \quad t = p + 2, ..., T. \tag{7}$$

The t-statistic for $\phi = 0$ is then denoted $ADF\text{-}GLS^{tb}(\tilde{\tau}, \bar{c})$. As is standard, we require that the lag truncation parameter, p, in (6) and (7) satisfies the following condition:

Assumption 2. *As $T \to \infty$, the lag truncation parameter p in (6) and (7) satisfies the condition that $1/p + p^2/T \to 0$.*

Remark 3: Perron and Rodríguez (2003) recommend the use of the modified Akaike Information Criterion (MAIC) of Ng and Perron (2001) for selecting p in in (6) and

(7) with an upper bound $p_{\max}$ that satisfies Assumption 2; see section 6 of Perron and Rodríguez (2003) for further details. $\square$

In Theorem 1 we now establish the asymptotic equivalence of $ADF\text{-}GLS^{tb}\left(\tilde{\tau}, \bar{c}\right)$ and $ADF\text{-}GLS^{tb}\left(\tau_0, \bar{c}\right)$.

Theorem 1 *Let y_t be generated according to (1) and (2), with $\gamma_0 \neq 0$, $\rho_T = 1 - c/T$, $0 \leq c < \infty$, and let Assumptions 1 and 2 hold. Then, provided $\tilde{\tau}$ is any $O_p\left(T^{-1}\right)$ consistent estimator of τ_0, it holds that $ADF\text{-}GLS^{tb}\left(\tilde{\tau}, \bar{c}\right) - ADF\text{-}GLS^{tb}\left(\tau_0, \bar{c}\right) \xrightarrow{p} 0$.*

This result with $c = 0$ shows that we can carry out the test $ADF\text{-}GLS^{tb}\left(\tilde{\tau}, \bar{c}\right)$ by using asymptotic critical values appropriate for $ADF\text{-}GLS^{tb}\left(\tau_0, \bar{c}\right)$. The asymptotic distribution of $ADF\text{-}GLS^{tb}\left(\tau_0, \bar{c}\right)$ is given in Theorem 1 of Perron and Rodríguez (2003) and the associated critical values are provided in their Table 1(b). For $c > 0$, the result confirms that $ADF\text{-}GLS^{tb}\left(\tilde{\tau}, \bar{c}\right)$ and $ADF\text{-}GLS^{tb}\left(\tau_0, \bar{c}\right)$ have identical asymptotic local alternative power functions.

Table 1 about here

The asymptotic Gaussian local power envelope for the testing problem considered in this section, where it is known that $\gamma_0 \neq 0$, is given in Perron and Rodríguez (2003,pp.7-8), who note that this function depends on the true break fraction τ_0. The value of c such that the Gaussian power envelope is at 0.50 is therefore expected to depend on both the significance level used and on τ_0. To investigate this further, Table 1 reports, for the grid of break fractions $\tau_0 \in \{0.15, 0.20, ..., 0.85\}$ and for the nominal 0.10, 0.05 and 0.01 significance levels, the corresponding values of $c = c_{\tau_0}$ for which the Gaussian power envelope is at 0.50. This was obtained by simulating the limit distribution of the point optimal invariant test of $c = 0$ against $c = c_{\tau_0}$, using the distributional result in equation (15) of Perron and Rodríguez (2003). Here we also give the corresponding asymptotic and finite sample critical values of $ADF\text{-}GLS^{tb}\left(\tau_0, c_{\tau_0}\right)$. The asymptotic critical values were obtained by simulating the limit distribution given in Theorem 1 of Perron and Rodríguez (2003). We approximate the Wiener processes in the limiting functionals using $NIID(0,1)$ random variates, and with the integrals approximated by normalized sums of 1000 steps. The finite sample critical values are also reported in Table 1. These were obtained by Monte Carlo simulation using the DGP (1) and (2) with $\gamma_0 = 0$, and setting $\alpha_0 = \beta_0 = 0$, $\varepsilon_t \sim NIID(0,1)$ and $u_1 = \varepsilon_1$. All simulations were based on 50,000 replications, using the `rndKMn` function of Gauss 7.0.

Because the value of c_{τ_0} varies with τ_0, as Table 1 demonstrates, in practice we would clearly like to calculate $ADF\text{-}GLS^{tb}\left(\tilde{\tau}, c_{\tilde{\tau}}\right)$, where $c_{\tilde{\tau}}$ denotes the value of c for which the asymptotic Gaussian local power envelope for a break fraction of $\tilde{\tau}$ is at 0.50. Now $c_{\tilde{\tau}}$ is clearly a random variable, but since $\tilde{\tau}$ is a consistent estimator of τ_0 it follows that $\lim_{T\to\infty} \Pr\left(c_{\tilde{\tau}} = c_{\tau_0}\right) = 1$ and, hence, that $\lim_{T\to\infty} \Pr\left(ADF\text{-}GLS^{tb}\left(\tilde{\tau}, c_{\tilde{\tau}}\right)\right.$ $= ADF\text{-}GLS^{tb}\left(\tilde{\tau}, c_{\tau_0}\right) = 1$. The associated critical value for $ADF\text{-}GLS^{tb}\left(\tilde{\tau}, c_{\tilde{\tau}}\right)$, denoted $v_{\tilde{\tau}}$ say, is also a random variable. But again the consistency of $\tilde{\tau}$ implies that

$\lim_{T\to\infty} \Pr(v_{\tilde{\tau}} = v_0) = 1$, where v_0 denotes the critical value for the true break fraction τ_0. It follows that the test statistic and critical value pairs $\left(ADF\text{-}GLS^{tb}(\tilde{\tau}, c_{\tilde{\tau}}), v_{\tilde{\tau}}\right)$ and $\left(ADF\text{-}GLS^{tb}(\tau_0, c_{\tau_0}), v_{\tau_0}\right)$ define asymptotically equivalent tests. As will subsequently be shown in section 5.1, $ADF\text{-}GLS^{tb}(\tilde{\tau}, c_{\tilde{\tau}})$ and, hence, $ADF\text{-}GLS^{tb}(\tau_0, c_{\tau_0})$ lie virtually on the asymptotic Gaussian local power envelope.

3.2 The Case where it is Unknown if a Break has Occurred

When no break in trend is present, i.e. when $\gamma_0 = 0$, such that τ_0 is not identified, it follows from Theorem 3.1 and Remark 2(a) of Nunes *et al.* (1995, p.741) that $\tilde{\tau}$ has a well defined asymptotic distribution with support equal to $\Lambda = [\tau_L, \tau_U]$; that is, $\tilde{\tau}$ does not converge in probability to a constant as in the case where $\gamma_0 \neq 0$; the same result holds for Perron and Rodríguez's (2003, p.6) QD estimator, $\hat{\tau}$. As a consequence the asymptotic null distribution of $ADF\text{-}GLS^{tb}(\tilde{\tau}, \bar{c})$ will differ according to whether a break in trend occurs or not.

In practice, where it will not be known if a trend break has in fact occurred or not, we are therefore faced with a choice when running unit root tests based on the estimated break fraction, $\tilde{\tau}$ (or, indeed, $\hat{\tau}$). We could, as in Perron and Rodríguez (2003), use conservative critical values corresponding to the case where it is assumed that no break is present, with a corresponding loss of efficiency in cases where a break is present. Alternatively, we could use (liberal) critical values which assume that a break is present, but run the risk of over-sizing, even asymptotically, in the unit root tests when a break is not present. Moreover, neither approach will deliver an efficient testing procedure in the no break case since here a redundant trend break regressor will always be included because of the behaviour of $\tilde{\tau}$ (and $\hat{\tau}$) noted above in this case; here the standard QD de-trended ADF-type test, including only an intercept and linear trend, of Elliott *et al.* (1996) is efficient.

The aforementioned problems with existing tests stem from the fact that the estimator of the break fraction does not converge to zero when no break in trend occurs. Clearly then, it would make considerable sense to use an estimator which, in the limit at least, places the estimated break fraction outside any range of break points considered to be feasible, that is outside of Λ, when no break occurs. At the same time, we would obviously want such an estimator to have consistency properties that are not inferior to those of $\tilde{\tau}$ in Theorem 1 when a break does actually occur. One way to achieve an estimator with these properties is to weight our first difference-based estimator $\tilde{\tau}$ by an auxiliary function of the data; that is, consider the modified estimator:

$$\bar{\tau} := (1 - \bar{\lambda})\tilde{\tau}. \tag{8}$$

The weight function $\bar{\lambda}$ in (8) needs to have the property that it converges to unity in such a way that $\bar{\tau}$ converges to zero at rate $O_p(T^{-1/2})$ when no break occurs, but converges to zero in such a way that $\bar{\tau}$ converges to to τ_0 at rate $O_p(T^{-1})$ when a break occurs, in each case irrespective of whether the unit root holds or not.

As we will subsequently show, a weight function which has this property is

$$\bar{\lambda} := \exp(-gT^{-1/2}W_T(\tilde{\tau})) \tag{9}$$

where g is some finite positive constant and, in the spirit of the work of Vogelsang (1998), $W_T(\tilde{\tau})$ denotes the (unscaled) Wald statistic for testing $\gamma_0 = 0$ in the partially-summed counterpart to regression equation (1), with $DT_t(\tau_0)$ replaced by $DT_t(\tilde{\tau})$. Supposing τ_0 to be known, $W_T(\tau_0)$ is constructed as follows. Calculate the residual sum of squares, $RSS_U(\tau_0)$, from the following partial sum regression, estimated using OLS,

$$\sum_{i=1}^{t} y_i = \alpha_0 t + \beta_0 \sum_{i=1}^{t} i + \gamma_0 \sum_{i=1}^{t} DT_i(\tau_0) + s_t, \quad t = 1, ..., T,$$

where $s_t := \sum_{i=1}^{t} u_i$, $t = 1, ..., T$, and calculate the residual sum of squares, RSS_R, from its restricted counterpart,

$$\sum_{i=1}^{t} y_i = \alpha_0 t + \beta_0 \sum_{i=1}^{t} i + s_t, \quad t = 1, ..., T,$$

and then calculate the (unscaled) Wald statistic

$$W_T(\tau_0) := \frac{RSS_R}{RSS_U(\tau_0)} - 1. \tag{10}$$

The statistic $W_T(\tilde{\tau})$ is then simply $W_T(\cdot)$ evaluated at $\tilde{\tau}$. Notice that, by definition, $0 \leq \bar{\lambda} \leq 1$, owing to the non-negativity of $W_T(\tilde{\tau})$.

In Lemma 2 we now establish the large sample behaviour of $W_T(\tau_0)$ and $W_T(\tilde{\tau})$ for both $\gamma = 0$ and $\gamma \neq 0$.

Lemma 2 *Let y_t be generated according to (1) and (2), with $\rho_T = 1 - c/T$, $c \geq 0$, and let Assumption 1 hold. Let $B_1(r) := \int_0^r B_0(s)\,ds$, where $B_0(r) := \int_0^r e^{-(r-s)c}dW(s)$ is a standard Ornstein-Uhlenbeck (OU) process, with $W(s)$ a standard Brownian motion on $[0,1]$, and let $Z_1(r) := (r, \frac{1}{2}r^2)'$, $Z_{2,\tau}(r) := 0 \vee \frac{1}{2}(r-\tau)^2$ and $Z_\tau(r) := (Z_1(r)', Z_{2,\tau}(r))'$. Finally, let S_1 and $S_{1,\tau}$ be the residual processes from a projection of B_1 on Z_1 and B_1 on Z_τ, respectively. Then,*

(i) If $\gamma_0 = 0$

$$W_T(\tau_0) \xrightarrow{d} \frac{\int_0^1 S_1(r)^2\,dr}{\int_0^1 S_{1,\tau_0}(r)^2\,dr} - 1$$

(ii) If $\gamma_0 \neq 0$

$$T^{-1}W_T(\tau_0) \xrightarrow{d} \gamma_0^2 \left(\int_0^1 S_{1,\tau_0}(r)^2\,dr \right)^{-1}$$

(iii) if $\gamma_0 = 0$, it holds that $W_T(\tilde{\tau}) = O_p(1)$
(iv) if $\gamma_0 \neq 0$, it holds that $T^{-1}W_T(\tilde{\tau}) = T^{-1}W_T(\tau_0) + o_p(1)$.

10

Lemma 2 shows that, regardless of which of $H_0 : c = 0$ or $H_1 : c > 0$ holds, both $W_T(\tau_0)$ and $W_T(\tilde{\tau})$ have the crucial property that they have well-defined (but not necessarily pivotal) large sample distributions when $\gamma_0 = 0$ but diverge at rate $O_p(T)$ when $\gamma_0 \neq 0$. The same can also be shown to hold in the stable autoregressive case. The results in Lemma 2 enable us to now establish in Lemma 3 the large sample properties of our new break fraction estimator, $\bar{\tau}$ of (8)-(9).

Lemma 3 *Let the conditions of Lemma 2 hold. (i) If $\gamma_0 = 0$, then $\bar{\tau} = O_p(T^{-1/2})$. (ii) If $\gamma_0 \neq 0$, then $T(\bar{\tau} - \tau_0) = O_p(1)$.*

The results of Lemma 3 (which also hold in the stable autoregressive case) imply that our new break fraction estimator, $\bar{\tau}$, 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. Consequently, and as required, $\bar{\tau}$ attains exactly the same rate of consistency as $\tilde{\tau}$ (cf. Lemma 1) when a break occurs, but avoids the problem of spuriously indicating a break when none is present.

Now consider the properties of QD detrended ADF-type unit root tests applied to (1) and (2) using our new break fraction estimator, $\bar{\tau}$. Based on the properties of $\bar{\tau}$ in Lemma 3, if $\bar{\tau} \geq \tau_L$ (notice that, by definition, $\bar{\tau}$ cannot exceed τ_U) then we take that as evidence of the presence of a trend break and correspondingly use $ADF\text{-}GLS^{tb}(\bar{\tau}, c_{\bar{\tau}})$ as our unit root test statistic, where $ADF\text{-}GLS^{tb}(\bar{\tau}, c_{\bar{\tau}})$ is identical to $ADF\text{-}GLS^{tb}(\tilde{\tau}, c_{\tilde{\tau}})$ of section 3.1, except that $\bar{\tau}$ replaces $\tilde{\tau}$, and where $c_{\bar{\tau}}$ denotes the value of c for which the asymptotic Gaussian local power envelope for a break fraction of $\bar{\tau}$ is at 0.50. In contrast, if $\bar{\tau} < \tau_L$ then we take that as evidence of the absence of a structural break and we then use a standard QD de-trended ADF-type test including only an intercept and linear trend (using $\bar{c} = 13.5$ for the QD transformation, as in Elliott *et al.*, 1996), denoted $ADF\text{-}GLS^t$, which is known to be an efficient test in this case. Our suggested unit root test statistic can therefore be written as:

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

In Theorem 2 we now establish the large sample behaviour of $t(\bar{\tau})$, demonstrating that, unlike existing procedures, it delivers an efficient test both where a trend break occurs and where one does not.

Theorem 2 *Let y_t be generated according to (1) and (2), with $\rho_T = 1 - c/T$, $0 \leq c < \infty$, and let Assumptions 1 and 2 hold.*
(i) If $\gamma_0 = 0$, then $t(\bar{\tau}) - ADF\text{-}GLS^t \xrightarrow{p} 0$.
(ii) If $\gamma_0 \neq 0$ then $t(\bar{\tau}) - ADF\text{-}GLS^{tb}(\tau_0, c_{\tau_0}) \xrightarrow{p} 0$.

The proof of Theorem 2 follows immediately from the results in Lemma 3. If $\gamma_0 = 0$ then Lemma 3(i) implies that $\lim_{T \to \infty} \Pr(\bar{\tau} < \tau_L) = 1$ and, hence, that $\lim_{T \to \infty} \Pr(t(\bar{\tau}) = ADF\text{-}GLS^{\,t}) = 1$ and we therefore consult the standard asymptotic

critical values that apply to $ADF\text{-}GLS^{\,t}$; see, Table 1 of Elliott et $al.$ (1996,p.825). Similarly, for $\gamma_0 \neq 0$, Lemma 3(ii) implies that $\lim_{T\to\infty} \Pr\left(\bar{\tau} \geq \tau_L\right) = 1$ and, hence, that $\lim_{T\to\infty} \Pr\left(t(\bar{\tau}) = ADF\text{-}GLS^{\,tb}\left(\bar{\tau}, c_{\bar{\tau}}\right)\right) = 1$. Moreover, because of the rate of convergence of $\bar{\tau}$ shown in Lemma 3(ii), Theorem 1 applies to show that $ADF\text{-}GLS^{\,tb}\left(\bar{\tau}, c_{\bar{\tau}}\right) - ADF\text{-}GLS^{\,tb}\left(\tau_0, c_{\tau_0}\right) \overset{p}{\to} 0$. Unlike procedures based on either the first difference or QD break fraction estimators, $\tilde{\tau}$ and $\hat{\tau}$, respectively, the practitioner is therefore not forced to make the choice between the conservative and liberal critical values discussed at the start of this section.

It follows from the results in Theorem 2 that where no trend break occurs, $t(\bar{\tau})$ has the same asymptotic local power function as the efficient $ADF\text{-}GLS^{\,t}$ test. Moreover, where a trend break occurs, $t(\bar{\tau})$ has the same asymptotic local power function as $ADF\text{-}GLS^{tb}\left(\tilde{\tau}, c_{\tilde{\tau}}\right)$ and, hence, $ADF\text{-}GLS^{tb}\left(\tau_0, c_{\tau_0}\right)$. Consequently in both the trend break and no break cases, $t(\bar{\tau})$ delivers a test which lies very close to the asymptotic Gaussian local power envelope.

4   An Approach based on Trend Break Pre-Testing

The approach adopted in using $t(\bar{\tau})$ in section 3.2 is not entirely dissimilar to using a pre-test for the presence of a trend break. The modified estimator $\bar{\tau}$ is effectively being used to form a decision rule as to whether a break has occurred or not, with a break being deemed to have occurred if $\bar{\tau} \geq \tau_L$. In such a case the $ADF\text{-}GLS^{tb}\left(\bar{\tau}, c_{\bar{\tau}}\right)$ statistic of Perron and Rodríguez (2003) is used, while if $\bar{\tau} < \tau_L$ the standard QD de-trended $ADF\text{-}GLS^{\,t}$ statistic of Elliott et $al.$ (1996) is used.

Other decision rules could clearly be used in an approach like this, in particular we might consider, as is done in Kim and Perron (2006) for the case of OLS de-trended unit root tests, using a formal statistical pre-test for the presence of a trend break. Like the weight function $\bar{\lambda}$ of (9), any such pre-test will need to possess certain large sample properties. Precisely, it needs to be based on a statistic which has a well-defined limiting distribution when $\gamma_0 = 0$, and it needs to be consistent when $\gamma_0 \neq 0$, with both of these properties holding regardless of whether the unit root holds or not. A number of trend break tests with these properties exist in the literature; see, $inter$ $alia$, Vogelsang and Perron (1998), Sayginsoy and Vogelsang (2004), Perron and Yabu (2006), and Harvey et $al.$ (2007b). The finite sample properties of a trend break pre-test will also be very important in practice because where a trend break does occur we want to be applying the $ADF\text{-}GLS^{tb}\left(\tilde{\tau}, c_{\tilde{\tau}}\right)$ test rather than $ADF\text{-}GLS^{\,t}$. Consequently, our trend break test needs to have good finite sample power. Of the available trend break tests, it is the test of Harvey et $al.$ (2007b) which displays the best overall power properties and so we shall focus on the use of that as a trend break pre-test in what follows.

The trend break test proposed by Harvey et $al.$ (2007b) rejects the null hypothesis

that $\gamma_0 = 0$ for large values of the statistic

$$t_\lambda := \lambda \left(\sup_{\tau \in \Lambda} |t_0(\tau)| \right) + m_\xi(1 - \lambda) \left(\sup_{\tau \in \Lambda} |t_1(\tau)| \right) \qquad (12)$$

where $t_0(\cdot)$ and $t_1(\cdot)$ are the OLS regression t-statistics for $\gamma_0 = 0$ in (1) and (3), respectively, in each case studentised using a long run variance estimator, and λ is the weight function

$$\lambda := \exp[-\{kS_0(\check{\tau})S_1(\tilde{\tau})\}^2]. \qquad (13)$$

In (13), k is a finite positive constant, and $S_0(\check{\tau})$ and $S_1(\tilde{\tau})$ are the stationarity test statistics of Kwiatkowski *et al.* (1992), calculated from the residuals $\{\hat{u}_t(\check{\tau})\}_{t=1}^T$ and $\{\tilde{v}_t(\tilde{\tau})\}_{t=2}^T$, respectively. Finally, as in Vogelsang (1998), m_ξ is a constant such that, for a significance level ξ, the asymptotic null critical value of of t_λ does not depend on whether the unit root holds or not. Selected critical values, together with the corresponding values of m_ξ, for the t_λ test are provided in Table 1 of Harvey *et al.* (2007b). Harvey *et al.* (2007b) show that when $\gamma_0 \neq 0$, t_λ is consistent at rate $O_p(T^{1/2})$ under the unit root (the same rate also holds under a near unit root) and at rate $O_p(T^{3/2})$ in the stable autoregressive case.

Using the trend break pre-test of Harvey *et al.* (2007b), we can therefore propose an alternative to $t(\bar{\tau})$ of (11), namely the statistic

$$t_P(\tilde{\tau}) := \begin{cases} ADF\text{-}GLS^{\,t} & \text{if } t_\lambda \text{ does not reject} \\ ADF\text{-}GLS^{\,tb}(\tilde{\tau}, c_{\tilde{\tau}}) & \text{if } t_\lambda \text{ rejects} \end{cases} \qquad (14)$$

If we choose a fixed (independent of sample size) significance level for the t_λ pre-test it should be clear that $t_P(\tilde{\tau})$ will not be asymptotically equivalent to $t(\bar{\tau})$. This occurs because even in large samples there will be a positive probability (given by the asymptotic significance level of t_λ) that the procedure will incorrectly select the $ADF\text{-}GLS^{\,tb}(\tilde{\tau}, c_{\tilde{\tau}})$ statistic when $\gamma_0 = 0$. This will therefore cause a degree of over-sizing (even asymptotically) in the test based on $t_P(\tilde{\tau})$, for the reason outlined at the start of section 3.2 in the context of running the $ADF\text{-}GLS^{tb}(\tilde{\tau}, c_{\tilde{\tau}})$ using liberal critical values. In order to avoid this problem and to obtain a procedure which is asymptotically equivalent to $t(\bar{\tau})$ we will need to follow the approach adopted by Kim and Perron (2006) and shrink the size of the pre-test, in our case t_λ, at a suitable rate in the sample size such that the test retains consistency when $\gamma_0 \neq 0$. Since t_λ diverges at rate $O_p(T^{1/2})$ when $\gamma_0 \neq 0$, this can clearly be achieved by defining the critical region of the pre-test to be of the form "*reject the null hypothesis $\gamma_0 = 0$ if $t_\lambda > cv_T$*" where $cv_T = aT^{1/2-d}$, for some $0 < d < 0.5$, and where a is a finite positive constant. For a given finite sample size, running the t_λ test at any conventional significance level is consistent with this decision rule.

5 Numerical Results

In this section we consider the performance of the $t(\bar{\tau})$ and $t_P(\tilde{\tau})$ tests proposed in this paper and assess the results relative to the performance of the two recommended

tests from Perron and Rodríguez (2003). The first of these is a Zivot and Andrews (1992)-type test which minimises the $ADF\text{-}GLS^{\,tb}\,(\tau, \bar{c})$ unit root statistic across all possible break dates

$$t_{ZA} := \inf_{\tau \in \Lambda} ADF\text{-}GLS^{\,tb}\,(\tau, \bar{c}).$$

The second comparator test is their conservative testing strategy, defined by

$$t_C(\hat{\tau}) := ADF\text{-}GLS^{\,tb}(\hat{\tau}, \bar{c})$$

where, as noted in footnote 2, $\hat{\tau}$ is the break date estimator obtained by maximising the absolute value of the t-ratio on the trend break dummy in the GLS regression (5) across $\tau \in \Lambda$. For both tests, Perron and Rodríguez (2003) recommend the use of $\bar{c} = -22.5$.

In section 5.1 we first examine the asymptotic power properties of the tests. Then in section 5.2 we turn to a comparison of the finite sample properties of the tests.

5.1 Asymptotic Results

In this section we simulate the asymptotic local to $I(1)$ power of the newly proposed tests, $t(\bar{\tau})$ and $t_P(\tilde{\tau})$, for both the no break and break cases. When no break exists, both $t(\bar{\tau})$ and $t_P(\tilde{\tau})$ are asymptotically equivalent to $ADF\text{-}GLS^{\,t}$, thus here we simulate the local asymptotic power curve for $ADF\text{-}GLS^{\,t}$, using the limiting functionals given in Elliott *et al.* (1996), for $\bar{c} = -13.5$. Otherwise, when a break is present, the $t(\bar{\tau})$ and $t_P(\tilde{\tau})$ tests are asymptotically equivalent to $ADF\text{-}GLS^{\,tb}\,(\tau_0, c_{\tau_0})$, the limit distribution of which is given in Theorem 1 of Perron and Rodríguez (2003). When a break exists, we also report the asymptotic Gaussian local power envelope; in the case of no break, the corresponding envelope is given in Elliott *et al.* (1996), where it is shown to be virtually indistinguishable from the power curve for $ADF\text{-}GLS^{\,t}$, thus we do not report it here.

The asymptotic distribution of t_{ZA} when no break is present was simulated using the limit expressions given in Perron and Rodríguez (2003). When a break exists, the t_{ZA} test does not have a pivotal limit distribution under the null or alternative, therefore we cannot simulate its asymptotic size or local power in any meaningful way. For $t_C(\hat{\tau})$, the asymptotic critical value is calculated under the assumption that no break exists, and was obtained by simulating the limit distribution for this test given in Perron and Rodríguez (2003). When a break does exist, the $t_C(\hat{\tau})$ statistic is asymptotically equivalent to the statistic $ADF\text{-}GLS^{\,tb}(\tau_0, \bar{c})$, allowing simulation of the test's asymptotic size and power, although note that it is the no break case (conservative) critical value that is still applied.

In Figures 1 and 2, we report results for $\gamma_0 = 0$ and $\gamma_0 \neq 0$, respectively, for $c \in \{0, 1, 2, ..., 50\}$. When a break exists, we consider three break fractions, $\tau_0 \in \{0.3, 0.5, 0.7\}$. As for the results in Table 1, we approximate the Wiener processes in the limiting functionals using $NIID(0, 1)$ random variates, and with the integrals approximated by normalized sums of 1000 steps and 50,000 replications, using the

14

rndKMn function of Gauss 7.0. All tests were calculated for the range of break fractions $\Lambda = [0.15, 0.85]$, and results are reported for the nominal 0.05 significance level.

<div align="center">**Figures $1 - 2$ about here**</div>

The results for the no break case in Figure 1 clearly demonstrate the loss in efficiency incurred in using either of the conservative t_{ZA} and $t_C(\hat{\tau})$ tests from Perron and Rodríguez (2003). These tests have virtually identical asymptotic local power functions which in both cases lie considerably inside the power functions of $t(\bar{\tau})$ and $t_P(\tilde{\tau})$; for example, for $c = 15$ the conservative tests both have power of approximately 0.25, while the $t(\bar{\tau})$ and $t_P(\tilde{\tau})$ tests both have power of around 0.60.

Turning to the results in Figure 2, for the case where a break in trend occurs, the conservative $t_C(\hat{\tau})$ test again has an asymptotic local power function which lies strictly inside the power function of $t(\bar{\tau})$ and $t_P(\tilde{\tau})$ which for the most part is virtually indistinguishable from the asymptotic Gaussian envelope. A comparison of the power function of $t_C(\hat{\tau})$ with those of $t(\bar{\tau})$ and $t_P(\tilde{\tau})$ highlights that, where a trend break occurs, the efficiency losses associated with using conservative critical values can be quite considerable; for example, when $\tau_0 = 0.7$, and $c = 25$, $t_C(\hat{\tau})$ has power of approximately 0.55 while $t(\bar{\tau})$ and $t_P(\tilde{\tau})$ have power of around 0.85. Interestingly, and consistent with the values of c_{τ_0} reported in Table 1, there is rather little variation in the shape of the asymptotic Gaussian local power envelope, and, hence, of the power functions $t(\bar{\tau})$ and $t_P(\tilde{\tau})$, across the three reported values of the break fraction τ_0.[5] The power function of $t_C(\hat{\tau})$ shows slightly greater dependence on τ_0, but even here this variation is not in any sense large.

5.2   Finite Sample Results

In this section we investigate the finite sample size and power properties of $t(\bar{\tau})$ and $t_P(\tilde{\tau})$, along with the comparator tests t_{ZA} and $t_C(\hat{\tau})$. Simulations are conducted for the DGP given by (1) and (2), with $\alpha_0 = \beta_0 = 0$ (without loss of generality), and $\varepsilon_t \sim NIID(0,1)$ with $u_1 = \varepsilon_1$. Sample sizes of $T = 150$ and 300 are considered and all simulations were computed for 20,000 replications, again using the rndKMn function of Gauss 7.0.

While Lemma 3 and Theorem 2, respectively, show that the constant g has no impact on the large sample properties of either $\bar{\tau}$ or $t(\bar{\tau})$ it will, however, influence the behaviour of $\bar{\tau}$ and $t(\bar{\tau})$ in finite samples. To gauge this effect, we allow for $g = 1.5$, 3 and 6 in the following simulations, denoting the resulting tests $t(\bar{\tau}, g)$. These values were selected on the basis of a number of preliminary simulation experiments and found to avoid extremes of either poor size or power that can arise with lower or higher values of the constant g.

---

[5]Correspondingly, although not reported here, we also found the local power functions of $t(\bar{\tau})$ and $t_P(\tilde{\tau})$ to be very insensitive to the value of the QD de-trending parameter used, being virtually indistinguishable from the asymptotic Gaussian power envelope for all but values of the QD de-trending parameter close to zero.

In the implementation of $t_P(\tilde{\tau})$, the pre-test t_λ was conducted with a significance level that shrank with the sample size. Specifically, we ran t_λ at the 0.05 level for $T = 150$ and at the 0.025 level for $T = 300$.[6] Following Harvey *et al.* (2007b) we set $k = 500$ in (13). We abstract from the issue of lag selection in the computation of all the tests, setting $p = 0$ in the ADF regressions.

In constructing both $t(\bar{\tau}, g)$ and $t_P(\tilde{\tau})$, we used a QD parameter of $\bar{c} = -13.5$ for $ADF\text{-}GLS^{\,t}$, while for $ADF\text{-}GLS^{\,tb}(\bar{\tau}, \bar{c})$, $\bar{c}$ was obtained by linear interpolation between the two nearest values of c_{τ_0} to either $\bar{\tau}$ for $t(\bar{\tau}, g)$ or $\tilde{\tau}$ for $t_P(\tilde{\tau})$, in the grid of values in Table 1. The corresponding critical values for $ADF\text{-}GLS^{\,tb}(\bar{\tau}, \bar{c})$ were also obtained by linear interpolation between the associated finite sample critical values in Table 1, and finite sample critical values were employed for $ADF\text{-}GLS^{\,t}$ (for the nominal 0.05 significance level, these are -2.96 for $T = 150$, and -2.92 for $T = 300$). Finite sample critical values were also used for the t_{ZA} and $t_C(\hat{\tau})$ tests; these were obtained by simulation using the DGP (1) and (2) with $\gamma_0 = 0$, using 50,000 Monte Carlo replications, and setting $\alpha_0 = \beta_0 = 0$, $\varepsilon_t \sim NIID(0, 1)$ and $u_1 = \varepsilon_1$ as before.

Our first set of finite sample simulations are designed to imitate the asymptotic cases in Figures 1 and 2. We therefore examine the autoregressive parameters $\rho_T = 1 - c/T$, $c \in \{0, 1, 2, ..., 50\}$. Results are reported for both cases where a break is not present in the DGP, and where a break of moderate magnitude exists; in this latter case, we set $\gamma_0 = 1$ (one standard deviation). Figure 3 reports results for the no break case, while Figure 4 reports results for a trend break occurring at $\tau_0 = 0.3$, 0.5 and 0.7.

Figures 3 − 4 about here

Consider first the results for the no break case in Figure 3. It can be seen that the three $t(\bar{\tau}, g)$ tests and the $t_P(\tilde{\tau})$ test all display a certain degree of over-sizing in finite samples, which is mitigated (albeit slowly) as the sample size is increased from $T = 150$ to $T = 300$. Since both approaches effectively employ a form of pre-test for the presence of a trend break and in this case there is no break in trend, this finite sample effect is to be expected (see the discussion of $t_P(\tilde{\tau})$ at the end of section 4). Among the $t(\bar{\tau}, g)$ tests, $t(\bar{\tau}, 1.5)$ yields sizes closest to the nominal 0.05 significance level. Here, both $t(\bar{\tau}, 1.5)$ and $t_P(\tilde{\tau})$ have size of 0.084 for $T = 150$ and, respectively, 0.077 and 0.071 for $T = 300$. For $t(\bar{\tau}, 3)$ and $t(\bar{\tau}, 6)$, the sizes change from 0.100 and 0.110, to 0.094 and 0.107, respectively, moving from $T = 150$ to $T = 300$. As regards power, the overall shapes of the power functions of all the tests are very similar to their asymptotic counterparts in Figure 1, particularly so for $T = 300$, with the $t(\bar{\tau}, g)$ and $t_P(\tilde{\tau})$ tests again showing almost identical power functions which clearly dominate those of the t_{ZA} and $t_C(\hat{\tau})$ tests.

Turning to the results for the trend break case in Figure 4, we see here that all of the $t(\bar{\tau}, g)$ tests, together with the $t_P(\tilde{\tau})$ test, have sizes very close to the nominal level,

---

[6]The critical values for the t_λ test using 15% trimming are 2.492 and 2.757 for the 0.05 and 0.025 significance levels, respectively. The corresponding b_α values also required to implement the t_λ test are, respectively, 0.849 and 0.867.

while, consistent with their large sample properties (see Figure 2), the t_{ZA} and $t_C(\hat{\tau})$ tests are somewhat under-sized. This effect is most pronounced for the $t_C(\hat{\tau})$ test and, as a consequence, the t_{ZA} test displays somewhat superior finite sample power than $t_C(\hat{\tau})$ throughout. The finite sample power curves for the $t(\bar{\tau}, g)$ and $t_P(\tilde{\tau})$ tests are again virtually indistinguishable from each other throughout Figure 4, and are generally very similar to their asymptotic counterparts in Figure 2 for values of c below around 20, but flatten off somewhat, relative to the results in Figure 2, for larger values of c. For the most part the $t(\bar{\tau}, g)$ and $t_P(\tilde{\tau})$ tests retain a considerable power advantage over the t_{ZA} and $t_C(\hat{\tau})$ tests, and increasingly so as the sample size grows, although for larger values of c the power functions do tend to cross; for example, when $\tau_0 = 0.5$ the power function of t_{ZA} crosses those of $t(\bar{\tau}, g)$ and $t_P(\tilde{\tau})$ at around $c = 30$ (where the power of both tests is about 0.84) when $T = 150$, while for $T = 300$ this crossing point has moved out to $c = 33$ (where power is about 0.90).

Our second set of finite sample simulations examine the size and power of the tests across breaks of varying magnitude, which is perhaps a more revealing analysis from a practitioner's standpoint. Here, we consider break magnitudes in the range $\gamma_0 \in \{0, 0.05, 0.10, ..., 2.00\}$ for the values $c = 0$, 10, 20 and 30 and, as above, break fractions $\tau_0 = 0.3$, 0.5 and 0.7.

Figures 5 − 8 about here

Figure 5 presents the size of the tests when $c = 0$. As might be expected given the previous simulations, sizes are highest for $\gamma_0 = 0$ and show a steady decrease as γ_0 increases. However, above about $\gamma_0 = 1$, the sizes remain virtually unchanged. The $t(\bar{\tau}, g)$ and $t_P(\tilde{\tau})$ tests are over-sized for small γ_0, but correctly sized for larger γ_0, while for t_{ZA} and $t_C(\hat{\tau})$, correct size obtains for small γ_0 and under-size for large γ_0. The respective sizes for $\gamma_0 = 0$ and $\gamma_0 = 1$ are of course identical to those at the origin in Figures 3 and 4.

Figures 6-8 display the power of the tests as functions of γ_0 for a given value of $c > 0$. For the $t(\bar{\tau}, g)$ and $t_P(\tilde{\tau})$ tests, highest power is achieved for $\gamma_0 = 0$ and while the power of these tests eventually becomes near constant for larger values of γ_0, each of these tests is prone to a perceptible dip in power for low, but non-zero, values of γ_0. The magnitude of this 'power trough' and how long (across γ_0) it is sustained depends on the particular test under consideration but also on the location of the breakpoint and value of c. Other things being equal, however, it is fairly clear that the phenomenon is generally most acute for the tests $t(\bar{\tau}, 1.5)$ and $t_P(\tilde{\tau})$, often taking their power well below that which is attainable with t_{ZA} and $t_C(\hat{\tau})$.

Considering $t_P(\tilde{\tau})$, it is fairly straightforward to see how this behaviour might arise. If γ_0 is zero, or indeed extremely small, the t_λ pre-test will not reject the null of no trend break and hence indicate that ADF-GLS^t is to be applied. This results in high unit root test power in this region as any unattended break in trend here is so small in magnitude that its neglect fails to induce non-rejection by ADF-GLS^t. For small values of γ_0, we enter a transitory region where, due to low power of t_λ, ADF-GLS^t is often wrongly being applied in place of ADF-$GLS^{tb}(\tilde{\tau}, c_{\tilde{\tau}})$. However, in these cases the

unattended break in trend is now large enough that $ADF\text{-}GLS^{\,t}$ will almost certainly never suggest a rejection of the unit root hypothesis. As a result, in this region of values of γ_0, the power of $t_P(\tilde{\tau})$ can be low. Once γ_0 becomes large enough, the power of t_λ is then high and so $ADF\text{-}GLS^{\,tb}(\tilde{\tau}, c_{\tilde{\tau}})$ is correctly applied, with unit root test power restored to the appropriate level.[7]

For the $t(\bar{\tau}, g)$ tests, the reasoning is somewhat similar. For some small γ_0, $W_T(\tilde{\tau})$ can be small and therefore $\bar{\lambda}$ is close to one. When τ_0 (and by implication $\tilde{\tau}$) is small, this results in $\bar{\tau} = (1 - \bar{\lambda})\tilde{\tau}$ often being below τ_L. Hence $t(\bar{\tau}, g)$ is set to be $ADF\text{-}GLS^{\,t}$, which does not reject the unit root hypothesis because a broken trend is present. For larger τ_0 (and $\tilde{\tau}$), $\bar{\tau}$ may no longer fall below τ_L, but is now sufficiently far from the true break date to result in non-rejections by $ADF\text{-}GLS^{\,tb}(\bar{\tau}, \bar{c})$. The extent to which these effects occur is controllable by the constant g: increasing g causes $\bar{\lambda}$ to move towards zero, thereby reducing the distance between $\bar{\tau}$ and $\tilde{\tau}$. This behaviour is quite evident from Figures 6-8; the depth and duration of the power troughs are much smaller for $t(\bar{\tau}, 6)$ than for $t(\bar{\tau}, 3)$, and these in turn are smaller than for $t(\bar{\tau}, 1.5)$. Of course, by increasing g, $t(\bar{\tau}, g)$ must also select $ADF\text{-}GLS^{\,t}$ less frequently when $\gamma_0 = 0$, so that power in this situation will be reduced. However, it is clear from Figures 6-8 that the extent of this power loss over the current settings for g - that is, comparing $t(\bar{\tau}, 6)$ to $t(\bar{\tau}, 1.5)$ when $\gamma_0 = 0$ - is only ever extremely modest.

On the basis of the foregoing finite sample simulation analysis we would recommend either $t(\bar{\tau}, 3)$ or $t(\bar{\tau}, 6)$ for empirical use. Both have reasonable size control and attractive power profiles relative to $t_P(\tilde{\tau})$, $t(\bar{\tau}, 1.5)$, $t_C(\hat{\tau})$ and t_{ZA}. If size is of particular concern, $t(\bar{\tau}, 3)$ should be favoured over $t(\bar{\tau}, 6)$, since its size across γ_0 does not exceed 0.10 for typical sample sizes. On the other hand, $t(\bar{\tau}, 6)$ does not have size above 0.11 for the sample sizes considered, and demonstrates a much superior ability to smooth over the power troughs that arise from small breaks in trend.

6   Conclusions

In this paper we have proposed new tests of the unit root null hypothesis, based on quasi difference de-trending, for the case where there is a possible one time change in the trend function occurring at an unknown point in the series. The first of these was based on using a new estimator of the break fraction which was shown, where a break occurs, to be consistent for the true break fraction at rate $O_p(T^{-1})$ both under the null hypothesis and local alternatives, but to not spuriously indicate the presence of a trend break where none exists. The second approach was based on first employing a pre-test for a break in trend, and allowing for a break in trend in the unit root regression only where this pre-test rejected. Our two proposed tests were shown to be asymptotically equivalent and, in contrast to extant tests in the literature, were shown to lie arbitrarily

---

[7]It is perhaps worth pointing out that it appears to be the problem of inappropriate selection of $ADF\text{-}GLS^{\,t}$ rather than simply inaccuracy of break point estimation in small samples that is mainly responsible for the $t_P(\tilde{\tau})$ power troughs, otherwise they should also be evident for $t_C(\hat{\tau})$.

close in large samples to the asymptotic Gaussian local limiting power envelope both where a break occurs and where a break does not occur. Asymptotic and finite sample evidence was reported which suggested that our two new tests generally outperformed other available tests in the literature.

To conclude we suggest two areas of possible further research. First, in this paper we have focused attention on classical testing methods, rather than Bayesian approaches to unit root testing or other model selection based methods, which might also be fruitfully explored. For further general discussion on these alternative approaches we direct the interested reader to, *inter alia*, Phillips (1991a,b), Phillips and Ploberger (1994) and Hansen (2007a,2007b). Second, our analysis allows for the possibility of a single break in trend. Extending the ideas in this paper to the case of multiple (deterministically occurring) trend breaks should be feasible using similar computational methods to those used in Carrion-i-Silvestre *et al.* (2007). For any more than two breaks, however, this procedure becomes problematic and a useful alternative is to consider the case where the breaks may be generated by an auxiliary stochastic component; important preliminary work in this area allowing for stochastic level shifts is considered in Cavaliere and Georgiev (2007).

Acknowledgements

A   Appendix

In what follows, we define $\|x\| = \sqrt{x'x}$ for any vector x and $\|A\| = \max_x \|Ax\| / \|x\|$ for a square matrix A.

A.1   Proof of Lemma 1

When $c = 0$, this follows immediately from Proposition 3 of Bai (1994) since $v_t = \Delta u_t$ satisfies Bai's Assumptions A and B. When $c > 0$ we need to generalise Bai's Proposition 1 so that the Hajek-Renyi equality applies to the first difference of a near-unit root process. Then the proof of Bai's Proposition 3 applies straightforwardly to conclude that $\tilde{\tau}$ is $O_p\left(T^{-1}\right)$ consistent when u_t is near-unit root. To do this, we substitute $u_t = \sum_{j=0}^{t-1} \rho_T^j \varepsilon_{t-j}$ into $v_t = \varepsilon_t - cT^{-1}u_{t-1}$ to obtain

$$v_t = \varepsilon_t - cT^{-1}\sum_{j=1}^{t-1} \rho_T^{t-j-1}\varepsilon_j.$$

19

Then some rearrangements give

$$\sum_{t=1}^{k} v_t = \sum_{t=1}^{k} a_t \varepsilon_t$$

for $k \leq T$, where $a_t = 1 - cT^{-1} \left(\sum_{j=0}^{k-t-1} \rho_T^j \right)$ for $t = 1, \ldots, k-1$ and $a_k = 1$. Note that $0 \leq a_t \leq 1$ because $\left| cT^{-1} \left(\sum_{j=0}^{k-t-1} \rho_T^j \right) \right| \leq cT^{-1} (k-t) \leq 1$. Now following Bai's proof of his Proposition 1 we apply a Beveridge and Nelson (1981) (BN) decomposition to ε_t:

$$\varepsilon_t = c(1) \eta_t - \Delta \varepsilon_t^*,$$

where $\varepsilon_t = \sum_{j=0}^{\infty} c_j^* \eta_{t-j}$ and $c_j^* = \sum_{i=j+1}^{\infty} c_i$, so that

$$\sum_{t=1}^{k} a_t \varepsilon_t = c(1) \sum_{t=1}^{k} a_t \eta_t - a_k \varepsilon_k^* - a_1 \varepsilon_0^* - cT^{-1} \sum_{t=1}^{k-1} \rho_T^{k-t-1} \varepsilon_t^*,$$

where we have used $a_{t+1} - a_t = cT^{-1} \rho_T^{k-t-1}$. Thus for $n \leq T$ we consider

$$\Pr \left(\max_{m \leq k \leq n} c_k \left| \sum_{t=1}^{k} v_t \right| > \alpha \right) \leq \Pr \left(\max_{m \leq k \leq n} c_k \left| c(1) \sum_{t=1}^{k} a_t \eta_t \right| > \alpha/4 \right)$$

$$+ \Pr \left(\max_{m \leq k \leq n} c_k \left| a_k \varepsilon_k^* \right| > \alpha/4 \right)$$

$$+ \Pr \left(c_m \left| \varepsilon_0^* \right| > \alpha/4 \right)$$

$$+ \Pr \left(\max_{m \leq k \leq n} c_k \left| cT^{-1} \sum_{t=1}^{k-1} \rho_T^{k-t-1} \varepsilon_t^* \right| > \alpha/4 \right) \text{ (A.1)}$$

By equation (2) of Hajek and Renyi (1955), the first probability in (A.1) satisfies

$$\Pr \left(\max_{m \leq k \leq n} c_k \left| c(1) \sum_{t=1}^{k} a_t \eta_t \right| > \alpha/4 \right) \leq \frac{16 c(1)^2}{\alpha^2} \left(c_m^2 \sum_{t=1}^{m} a_t^2 + \sum_{k=m+1}^{n} c_k^2 a_k^2 \right)$$

$$\leq \frac{16 c(1)^2}{\alpha^2} \left(m c_m^2 + \sum_{k=m+1}^{n} c_k^2 \right), \quad \text{(A.2)}$$

the second line following because $0 \leq a_k \leq 1$. The second probability in (A.1) satisfies

$$\Pr \left(\max_{m \leq k \leq n} c_k \left| a_k \varepsilon_k^* \right| > \alpha/4 \right) \leq \sum_{k=m}^{n} \Pr \left(c_k \left| a_k \varepsilon_k^* \right| > \alpha/4 \right) \leq \frac{16 \sigma_*^2}{\alpha^2} \left(c_m^2 + \sum_{k=m+1}^{n} c_k^2 \right)$$

$$\text{(A.3)}$$

by Chebyshev's inequality, where $\sigma_*^2 = \sum_{j=0}^{\infty} c_j^{*2}$ which exists under Assumption 1. Again by Chebyshev's inequality, the third probability in (A.1) satisfies

$$\Pr \left(c_m \left| \varepsilon_0^* \right| > \alpha/4 \right) \leq \frac{16 \sigma_*^2}{\alpha^2} c_m^2. \quad \text{(A.4)}$$

The fourth probability in (A.1) satisfies

$$\Pr\left(\max_{m\leq k\leq n} c_k \left|cT^{-1}\sum_{t=1}^{k-1}\rho_T^{k-t-1}\varepsilon_t^*\right| > \alpha/4\right) \leq \sum_{k=m}^{n}\Pr\left(c_k\left|cT^{-1}\sum_{t=1}^{k-1}\rho_T^{k-t-1}\varepsilon_t^*\right| > \alpha/4\right)$$

$$\leq \frac{16}{\alpha^2}\sum_{k=m}^{n}c_k^2\mathsf{var}\left(cT^{-1}\sum_{t=1}^{k-1}\rho_T^{k-t-1}\varepsilon_t^*\right)$$

$$\leq \frac{16c^2\omega_*^2}{\alpha^2 T}\sum_{k=m}^{n}c_k^2, \tag{A.5}$$

by Chebyshev's inequality, where $\omega_*^2 = \sum_{j=-\infty}^{\infty}\left|\gamma_j^*\right|$ and $\gamma_j^* = E\left(\varepsilon_t^*\varepsilon_{t-j}^*\right)$ (and ω_*^2 exists under Assumption 1), since

$$\mathsf{var}\left(cT^{-1}\sum_{t=1}^{k-1}\rho_T^{k-t-1}\varepsilon_t^*\right) = c^2 T^{-2}\sum_{t=1}^{k-1}\sum_{s=1}^{k-1}\rho_T^{2k-t-s-2}\gamma_{t-s}^*$$

$$\leq c^2 T^{-2}\sum_{t=1}^{k-1}\sum_{s=1}^{k-1}\left|\gamma_{t-s}^*\right| \leq c^2 T^{-2}k\omega_*^2 \leq c^2 T^{-1}\omega_*^2.$$

Thus combining (A.2)–(A.5) in (A.1) leads to

$$\Pr\left(\max_{m\leq k\leq n} c_k\left|\sum_{t=1}^{k}v_t\right| > \alpha\right) \leq \frac{16}{\alpha^2}\left(c\left(1\right)^2 + 2\sigma_*^2 + c^2\omega_*^2\right)\left(mc_m^2 + \sum_{k=m+1}^{n}c_k^2\right),$$

which provides the required generalisation of Bai's Proposition 1.

A.2   Proof of Theorem 1

It will be convenient to represent the models in stacked matrix form as

$$y = X\theta_0 + u \tag{A.6}$$

and

$$y_{\bar{c}} = X_{\bar{c}}\theta_0 + u_{\bar{c}}, \tag{A.7}$$

where $\theta_0 = (\alpha_0, \beta_0, \gamma_0)'$, $X = (X_1(\tau_0), \ldots, X_T(\tau_0))'$, and so on. The OLS estimator of θ_0 based on (A.7) taking τ_0 as known is denoted $\hat{\theta}_{\bar{c}}$ and the resulting residuals from (A.6) are $\hat{u} = y - X\hat{\theta}_{\bar{c}}$. The ADF test statistic for $\hat{u}$ can be represented

$$ADF\text{-}GLS^{tb}\left(\tau_0, \bar{c}\right) = \frac{\hat{u}_{-1}'\hat{P}\Delta\hat{u}}{\hat{\sigma}\left(\hat{u}_{-1}'\hat{P}\hat{u}_{-1}\right)^{1/2}}, \tag{A.8}$$

with $\Delta\hat{u} = (\Delta\hat{u}_t)_{t=p+2}^T$, $\hat{u}_{-1} = (\hat{u}_{t-1})_{t=p+2}^T$, $\hat{U}_p = (\Delta\hat{u}_{t-1}, \ldots, \Delta\hat{u}_{t-p})_{t=p+2}^T$, $\hat{P} = I_{T-p-1} - \hat{U}_p \left(\hat{U}_p'\hat{U}_p\right)^{-1} \hat{U}_p'$ and

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

When $\tilde{\tau}$ replaces 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 resulting levels residuals are $\tilde{u} = y - \tilde{X}\tilde{\theta}_{\bar{c}}$. The statistic $ADF\text{-}GLS^{tb}(\tilde{\tau})$ is defined as in (A.8) but with $\hat{u}$ and $\hat{\sigma}$ replaced by $\tilde{u}$ and $\tilde{\sigma}$. We will show that

$$ADF\text{-}GLS^{tb}(\tilde{\tau}, \bar{c}) - ADF\text{-}GLS^{tb}(\tau_0, \bar{c}) \xrightarrow{p} 0. \tag{A.9}$$

We first provide some preliminary results. The first concerns the difference between the de-trending coefficients $\tilde{\theta}_{\bar{c}}$ and $\hat{\theta}_{\bar{c}}$.

Lemma 4 *Define $D_T = \mathsf{diag}\left(1, T^{1/2}, T^{1/2}\right)$. Then*

$$D_T\left(\hat{\theta}_{\bar{c}} - \theta_0\right) \xrightarrow{d} \left(\begin{array}{c} \varepsilon_1 \\ \left(\int_0^1 H_{\bar{c},\tau_0}(s) H_{\bar{c},\tau_0}(s)' ds\right)^{-1}\int_0^1 H_{\bar{c},\tau_0}(s)\left(dB_c(s) + \bar{c}B_c(s) ds\right) \end{array}\right), \tag{A.10}$$

where B_c is an OU process with long run variance ω_ε^2 and

$$H_{\bar{c},\tau_0}(s) = \left(\begin{array}{c} 1 + \bar{c}s \\ 1(s > \tau_0)(1 + \bar{c}(s - \tau_0)) \end{array}\right),$$

and $\tilde{\theta}_{\bar{c}}$ is asymptotically equivalent to $\hat{\theta}_{\bar{c}}$ in the sense that

$$D_T\left(\tilde{\theta}_{\bar{c}} - \hat{\theta}_{\bar{c}}\right) = O_p\left(T^{-1/2}\right). \tag{A.11}$$

The next lemma concerns the differences between individual sample statistics involving $\tilde{u}_t$ and $\hat{u}_t$.

Lemma 5 *(i) $T^{-1/2}\left(\tilde{u}_{\lfloor Ts\rfloor} - \hat{u}_{\lfloor Ts\rfloor}\right) = O_p\left(T^{-1/2}\right)$ uniformly for $s \in [0,1]$*
(ii) $T^{-1}\sum_{t=p+1}^T \Delta\tilde{u}_{t-i}\Delta\tilde{u}_{t-j} - T^{-1}\sum_{t=p+1}^T \Delta\hat{u}_{t-i}\Delta\hat{u}_{t-j} = O_p\left(T^{-1}\right)$ uniformly in $i, j = 1, \ldots p$
(iii) $T^{-1}\sum_{t=p+1}^T \tilde{u}_{t-1}\Delta\tilde{u}_{t-i} - T^{-1}\sum_{t=p+1}^T \hat{u}_{t-1}\Delta\hat{u}_{t-i} = O_p\left(T^{-1/2}\right)$ uniformly in $i = 1, \ldots p$.

The following orders of magnitude follow from Lemma 3.2 of Chang and Park (2002)[8]

---

[8]As noted in their Remark 3.1, their results continue to hold when applied to models with de-trending.

$$\left\| \left(T^{-1}\hat{U}_p'\hat{U}_p \right)^{-1} \right\| = O_p\left(1\right) \tag{A.12}$$

$$\left\| T^{-1}\hat{u}_{-1}'\hat{U}_p \right\| = O_p\left(p^{1/2}\right) \tag{A.13}$$

The next Lemma shows how the corresponding matrices behave when computed using $\tilde{u}_t$.

Lemma 6 *Corresponding to (A.12), (A.13), we can show*

(i) $\left\| \left(T^{-1}\tilde{U}_p'\tilde{U}_p \right)^{-1} \right\| = O_p\left(1\right),$

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

Also

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

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

(v) $\left\| \left(T^{-1}\tilde{U}_p'\tilde{U}_p \right)^{-1} - \left(T^{-1}\hat{U}_p'\hat{U}_p \right)^{-1} \right\| = O_p\left(pT^{-1}\right),$

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

From Lemma 5(iii) we have

$$\frac{\tilde{u}_{-1}'\Delta\tilde{u}}{T} - \frac{\hat{u}_{-1}'\Delta\hat{u}}{T} \xrightarrow{p} 0,$$

and from (A.12) and the results of Lemma 6 we find

$$\left| \frac{\tilde{u}_{-1}'\tilde{U}_p}{T} \left(\frac{\tilde{U}_p'\tilde{U}_p}{T} \right)^{-1} \frac{\tilde{U}_p'\Delta\tilde{u}}{T} - \frac{\hat{u}_{-1}'\hat{U}_p}{T} \left(\frac{\hat{U}_p'\hat{U}_p}{T} \right)^{-1} \frac{\hat{U}_p'\Delta\hat{u}}{T} \right|$$

$$\leq \left\| \frac{\tilde{u}_{-1}'\tilde{U}_p}{T} \right\| \left\| \left(\frac{\tilde{U}_p'\tilde{U}_p}{T} \right)^{-1} \right\| \left\| \frac{\tilde{U}_p'\Delta\tilde{u}}{T} - \frac{\hat{U}_p'\Delta\hat{u}}{T} \right\|$$

$$+ \left\| \frac{\tilde{u}_{-1}'\tilde{U}_p}{T} \right\| \left\| \left(\frac{\tilde{U}_p'\tilde{U}_p}{T} \right)^{-1} - \left(\frac{\hat{U}_p'\hat{U}_p}{T} \right)^{-1} \right\| \left\| \frac{\hat{U}_p'\Delta\hat{u}}{T} \right\|$$

$$+ \left\| \frac{\tilde{u}_{-1}'\tilde{U}_p}{T} - \frac{\hat{u}_{-1}'\hat{U}_p}{T} \right\| \left\| \left(\frac{\hat{U}_p'\hat{U}_p}{T} \right)^{-1} \right\| \left\| \frac{\hat{U}_p'\Delta\hat{u}}{T} \right\|$$

$$= O_p\left(pT^{-1/2}\right),$$

from which we can conclude

$$\frac{\tilde{u}_{-1}'\tilde{P}\Delta\tilde{u}}{T} - \frac{\hat{u}_{-1}'\hat{P}\Delta\hat{u}}{T} \xrightarrow{p} 0. \tag{A.14}$$

23

From Lemma 5 we have

$$\frac{\tilde{u}_{-1}'\tilde{u}_{-1}}{T^2} - \frac{\hat{u}_{-1}'\hat{u}_{-1}}{T^2} \overset{p}{\to} 0,$$

and from (A.12) and the results of Lemma 6 we find

$$T^{-1}\left|\frac{\tilde{u}_{-1}'\tilde{U}_p}{T}\left(\frac{\tilde{U}_p'\tilde{U}_p}{T}\right)^{-1}\frac{\tilde{U}_p'\tilde{u}_{-1}}{T} - \frac{\hat{u}_{-1}'\hat{U}_p}{T}\left(\frac{\hat{U}_p'\hat{U}_p}{T}\right)^{-1}\frac{\hat{U}_p'\hat{u}_{-1}}{T}\right|$$

$$\leq \quad T^{-1}\left\|\frac{\tilde{u}_{-1}'\tilde{U}_p}{T}\right\|\left\|\left(\frac{\tilde{U}_p'\tilde{U}_p}{T}\right)^{-1}\right\|\left\|\frac{\tilde{U}_p'\tilde{u}_{-1}}{T} - \frac{\hat{U}_p'\hat{u}_{-1}}{T}\right\|$$

$$+T^{-1}\left\|\frac{\tilde{u}_{-1}'\tilde{U}_p}{T}\right\|\left\|\left(\frac{\tilde{U}_p'\tilde{U}_p}{T}\right)^{-1} - \left(\frac{\hat{U}_p'\hat{U}_p}{T}\right)^{-1}\right\|\left\|\frac{\hat{U}_p'\hat{u}_{-1}}{T}\right\|$$

$$+T^{-1}\left\|\frac{\tilde{u}_{-1}'\tilde{U}_p}{T} - \frac{\hat{u}_{-1}'\hat{U}_p}{T}\right\|\left\|\left(\frac{\hat{U}_p'\hat{U}_p}{T}\right)^{-1}\right\|\left\|\frac{\hat{U}_p'\hat{u}_{-1}}{T}\right\|$$

$$= \quad O_p\left(pT^{-3/2}\right),$$

such that

$$\frac{\tilde{u}_{-1}'\tilde{P}\tilde{u}_{-1}}{T^2} - \frac{\hat{u}_{-1}'\hat{P}\hat{u}_{-1}}{T^2} \overset{p}{\to} 0. \tag{A.15}$$

From Lemma 5(ii) we have

$$\frac{\Delta\tilde{u}'\Delta\tilde{u}}{T} - \frac{\Delta\hat{u}'\Delta\hat{u}}{T} \overset{p}{\to} 0,$$

and from (A.12) and the results of Lemma 6 we find

$$\left|\frac{\Delta\tilde{u}'\tilde{U}_p}{T}\left(\frac{\tilde{U}_p'\tilde{U}_p}{T}\right)^{-1}\frac{\tilde{U}_p'\Delta\tilde{u}}{T} - \frac{\Delta\hat{u}'\hat{U}_p}{T}\left(\frac{\hat{U}_p'\hat{U}_p}{T}\right)^{-1}\frac{\hat{U}_p'\Delta\hat{u}}{T}\right|$$

$$\leq \quad \left\|\frac{\Delta\tilde{u}'\tilde{U}_p}{T}\right\|\left\|\left(\frac{\tilde{U}_p'\tilde{U}_p}{T}\right)^{-1}\right\|\left\|\frac{\tilde{U}_p'\Delta\tilde{u}}{T} - \frac{\hat{U}_p'\Delta\hat{u}}{T}\right\|$$

$$+\left\|\frac{\Delta\tilde{u}'\tilde{U}_p}{T}\right\|\left\|\left(\frac{\tilde{U}_p'\tilde{U}_p}{T}\right)^{-1} - \left(\frac{\hat{U}_p'\hat{U}_p}{T}\right)^{-1}\right\|\left\|\frac{\hat{U}_p'\Delta\hat{u}}{T}\right\|$$

$$+\left\|\frac{\Delta\tilde{u}'\tilde{U}_p}{T} - \frac{\Delta\hat{u}'\hat{U}_p}{T}\right\|\left\|\left(\frac{\hat{U}_p'\hat{U}_p}{T}\right)^{-1}\right\|\left\|\frac{\hat{U}_p'\Delta\hat{u}}{T}\right\|$$

$$= \quad O_p\left(p^2T^{-1}\right),$$

24

hence

$$\frac{\Delta\tilde{u}'\tilde{P}\Delta\tilde{u}}{T} - \frac{\Delta\hat{u}'\hat{P}\Delta\hat{u}}{T} \xrightarrow{p} 0. \tag{A.16}$$

From (A.14), (A.15) and (A.16), we conclude that (A.9) holds.

Proof of Lemma 4.

To prove (A.10) we first note the quasi-difference of $DT_t(\tau_0)$ can be written

$$
\begin{aligned}
DT_{\bar{c},t}(\tau_0) &= \Delta DT_t(\tau_0) + \bar{c}T^{-1}DT_{t-1}(\tau_0) \\
&= 1\left(t > \lfloor\tau_0 T\rfloor\right)\left(\left(1 + \bar{c}\left(t - 1 - \lfloor\tau_0 T\rfloor\right)/T\right)\right)
\end{aligned}
$$

Using

$$
X_{\bar{c},1}(\tau_0) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad X_{\bar{c},t}(\tau_0) = \begin{pmatrix} \bar{c}/T \\ 1 + \bar{c}(t-1)/T \\ 1\left(t > \lfloor\tau_0 T\rfloor\right)\left(\left(1 + \bar{c}\left(t - 1 - \lfloor\tau_0 T\rfloor\right)/T\right)\right) \end{pmatrix}
$$

we find

$$
\begin{aligned}
D_T^{-1}X_{\bar{c}}'X_{\bar{c}}D_T^{-1} &= \sum_{t=1}^{T} D_T^{-1}X_{\bar{c},t}(\tau_0)X_{\bar{c},t}(\tau_0)'D_T^{-1} \\
&\to \begin{pmatrix} 1 & 0 \\ 0 & \int_0^1 H_{\bar{c},\tau_0}(s)H_{\bar{c},\tau_0}(s)'\,ds \end{pmatrix}.
\end{aligned} \tag{A.17}
$$

Similarly

$$
D_T^{-1}X_{\bar{c}}'u_{\bar{c}} \xrightarrow{d} \begin{pmatrix} \varepsilon_1 \\ \int_0^1 H_{\bar{c},\tau_0}(s)\left((dB_c(s) + \bar{c}B_c(s)\,ds)\right) \end{pmatrix}
$$

so that

$$
D_T\left(\hat{\theta}_{\bar{c}} - \theta_0\right) \xrightarrow{d} \begin{pmatrix} \varepsilon_1 \\ \left(\int_0^1 H_{\bar{c},\tau_0}(s)H_{\bar{c},\tau_0}(s)'\,ds\right)^{-1}\int_0^1 H_{\bar{c},\tau_0}(s)\left((dB_c(s) + \bar{c}B_c(s)\,ds)\right) \end{pmatrix}.
$$

Next to show (A.11) we subtract

$$X_{\bar{c}}'X_{\bar{c}}\hat{\theta}_{\bar{c}} = X_{\bar{c}}'X_{\bar{c}}\theta_0 + X_{\bar{c}}'u_{\bar{c}}$$

from

$$\tilde{X}_{\bar{c}}'\tilde{X}_{\bar{c}}\tilde{\theta}_{\bar{c}} = \tilde{X}_{\bar{c}}'X_{\bar{c}}\theta_0 + \tilde{X}_{\bar{c}}'u_{\bar{c}}$$

and rearrange to give

$$
\begin{aligned}
D_T\left(\tilde{\theta}_{\bar{c}} - \hat{\theta}_{\bar{c}}\right) &= \left(D_T^{-1}\tilde{X}_{\bar{c}}'\tilde{X}_{\bar{c}}D_T^{-1}\right)^{-1}\left[-\left(D_T^{-1}\tilde{X}_{\bar{c}}'\tilde{X}_{\bar{c}}D_T^{-1} - D_T^{-1}X_{\bar{c}}'X_{\bar{c}}D_T^{-1}\right)D_T\left(\hat{\theta}_{\bar{c}} - \theta_0\right)\right. \\
&\quad \left. - \left(D_T^{-1}\tilde{X}_{\bar{c}}'\tilde{X}_{\bar{c}} - D_T^{-1}\tilde{X}_{\bar{c}}'X_{\bar{c}}\right)\theta_0 + D_T^{-1}\tilde{X}_{\bar{c}}'u_{\bar{c}} - D_T^{-1}X_{\bar{c}}'u_{\bar{c}}\right].
\end{aligned}
$$

25

We will prove (A.11) by showing

$$D_T^{-1}\tilde{X}_{\bar{c}}'u_{\bar{c}} - D_T^{-1}X_{\bar{c}}'u_{\bar{c}} = O_p\left(T^{-1/2}\right), \tag{A.18}$$

$$\left(D_T^{-1}\tilde{X}_{\bar{c}}'\tilde{X}_{\bar{c}} - D_T^{-1}\tilde{X}_{\bar{c}}'X_{\bar{c}}\right)\theta_0 = O_p\left(T^{-1/2}\right), \tag{A.19}$$

$$D_T^{-1}\tilde{X}_{\bar{c}}'\tilde{X}_{\bar{c}}D_T^{-1} - D_T^{-1}X_{\bar{c}}'X_{\bar{c}}D_T^{-1} = O_p\left(T^{-1/2}\right), \tag{A.20}$$

Note that this last line and the invertibility of (A.17) shows that $\left(D_T^{-1}\tilde{X}_{\bar{c}}'\tilde{X}_{\bar{c}}D_T^{-1}\right)^{-1}$ exists in the limit.

To show (A.18), we write

$$D_T^{-1}\tilde{X}_{\bar{c}}'u_{\bar{c}} - D_T^{-1}X_{\bar{c}}'u_{\bar{c}} = \begin{pmatrix} 0 \\ 0 \\ T^{-1/2}\sum_{t=1}^{T}\left(DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right)\right)u_{\bar{c},t} \end{pmatrix}.$$

Without loss of generality we will proceed as if $\tilde{\tau} < \tau_0$. Then

$$DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right) = \begin{cases} 0, & t \le \lfloor\tilde{\tau}T\rfloor \\ 1 + \bar{c}\left(t - 1 - \lfloor\tilde{\tau}T\rfloor\right)/T, & \lfloor\tilde{\tau}T\rfloor < t \le \lfloor\tau_0 T\rfloor \\ \bar{c}\left(\lfloor\tau_0 T\rfloor - \lfloor\tilde{\tau}T\rfloor\right)/T, & \lfloor\tau_0 T\rfloor < t \le T. \end{cases}$$

Note that the second component of this expression disappears if $\lfloor\tilde{\tau}T\rfloor = \lfloor\tau_0 T\rfloor$. Now

$$T^{-1/2}\sum_{t=1}^{T}\left(DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right)\right)u_{\bar{c},t}$$

$$= T^{-1/2}\sum_{\lfloor\tilde{\tau}T\rfloor < t \le \lfloor\tau_0 T\rfloor}\left(1 + \bar{c}\left(t - 1 - \lfloor\tilde{\tau}T\rfloor\right)/T\right)u_{\bar{c},t} + \bar{c}\frac{\lfloor\tau_0 T\rfloor - \lfloor\tilde{\tau}T\rfloor}{T}T^{-1/2}\sum_{\lfloor\tau_0 T\rfloor < t \le T}u_{\bar{c},t},$$

in which the second term is $O_p\left(T^{-1}\right)$ since

$$\frac{\lfloor\tau_0 T\rfloor - \lfloor\tilde{\tau}T\rfloor}{T} = \left(\tilde{\tau} - \tau_0\right) + \frac{\left(\lfloor\tau_0 T\rfloor - \tau_0 T\right)}{T} + \frac{\left(\tilde{\tau}T - \lfloor\tilde{\tau}T\rfloor\right)}{T} = O_p\left(T^{-1}\right), \tag{A.21}$$

while the first term has order $O_p\left(T^{-1/2}\left(\lfloor\tau_0 T\rfloor - \lfloor\tilde{\tau}T\rfloor\right)^{1/2}\right) = O_p\left(T^{-1/2}\right)$ also by (A.21).

To show (A.19), we write

$$\left(D_T^{-1}\tilde{X}_{\bar{c}}'\tilde{X}_{\bar{c}} - D_T^{-1}\tilde{X}_{\bar{c}}'X_{\bar{c}}\right)\theta_0$$

$$= D_T^{-1}\tilde{X}_{\bar{c}}'\left(\tilde{X}_{\bar{c}} - X_{\bar{c}}\right)\theta_0$$

$$= \sum_{t=1}^{T}D_T^{-1}X_{\bar{c},t}\left(\tilde{\tau}\right)\left(DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right)\right)\gamma_0$$

$$= \begin{pmatrix} \bar{c}T^{-1}\sum_{t=2}^{T}\left(DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right)\right)\gamma_0 \\ T^{-1/2}\sum_{t=2}^{T}\left(1 + \bar{c}\left(t - 1\right)/T\right)\left(DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right)\right)\gamma_0 \\ T^{-1/2}\sum_{\lfloor\tilde{\tau}T\rfloor < t \le T}\left(\left(1 + \bar{c}\left(t - 1 - \lfloor\tilde{\tau}T\rfloor\right)/T\right)\right)\left(DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right)\right)\gamma_0 \end{pmatrix}.$$

26

The first term disappears because

$$T^{-1} \sum_{t=2}^{T} \left(DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right) \right)$$

$$= T^{-1} \sum_{\lfloor \tilde{\tau}T \rfloor < t \le \lfloor \tau_0 T \rfloor} 1 + \bar{c}\left(t - 1 - \lfloor \tilde{\tau}T \rfloor\right)/T + T^{-1} \sum_{\lfloor \tau_0 T \rfloor < t \le T} \bar{c}\left(\lfloor \tau_0 T \rfloor - \lfloor \tilde{\tau}T \rfloor\right)/T$$

$$\le \frac{\lfloor \tau_0 T \rfloor - \lfloor \tilde{\tau}T \rfloor}{T}\left(1 + \bar{c}\frac{\lfloor \tau_0 T \rfloor - \lfloor \tilde{\tau}T \rfloor}{T}\right) + \bar{c}\frac{T - \lfloor \tau_0 T \rfloor}{T}\frac{\lfloor \tau_0 T \rfloor - \lfloor \tilde{\tau}T \rfloor}{T}$$

$$= O_p\left(T^{-1}\right),$$

so the second term then immediately follows using

$$T^{-1/2} \sum_{t=2}^{T} \left(1 + \bar{c}\left(t - 1\right)/T\right)\left(DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right)\right)$$

$$\le (1 + \bar{c}) T^{-1/2} \sum_{t=2}^{T} \left(DT_{\bar{c},t}\left(\tilde{\tau}\right) - DT_{\bar{c},t}\left(\tau_0\right)\right)$$

$$= O_p\left(T^{-1/2}\right)$$

and similarly for the third term.

To show (A.20), we simply write

$$D_T^{-1}\tilde{X}'_{\bar{c}}\tilde{X}_{\bar{c}}D_T^{-1} - D_T^{-1}X'_{\bar{c}}X_{\bar{c}}D_T^{-1} = \left(D_T^{-1}\tilde{X}'_{\bar{c}}\tilde{X}_{\bar{c}}D_T^{-1} - D_T^{-1}\tilde{X}'_{\bar{c}}X_{\bar{c}}D_T^{-1}\right)$$

$$+ \left(D_T^{-1}\tilde{X}'_{\bar{c}}X_{\bar{c}}D_T^{-1} - D_T^{-1}X'_{\bar{c}}X_{\bar{c}}D_T^{-1}\right),$$

and it is then easy to see these terms disappear using the steps used to show (A.19).

Proof of Lemma 5

(i) With the notation $\hat{\theta}_{\bar{c}} = \left(\hat{\alpha}_{\bar{c}}, \hat{\beta}_{\bar{c}}, \hat{\gamma}_{\bar{c}}\right)'$ and $\tilde{\theta}_{\bar{c}} = \left(\tilde{\alpha}_{\bar{c}}, \tilde{\beta}_{\bar{c}}, \tilde{\gamma}_{\bar{c}}\right)'$, we can write

$$\hat{u}_t = u_t - \left(\hat{\alpha}_{\bar{c}} - \alpha_0\right) - \left(\hat{\beta}_{\bar{c}} - \beta_0\right)t - \left(\hat{\gamma}_{\bar{c}} - \gamma_0\right)DT_t\left(\tau_0\right)$$

and

$$\tilde{u}_t = u_t - \left(\tilde{\alpha}_{\bar{c}} - \alpha_0\right) - \left(\tilde{\beta}_{\bar{c}} - \beta_0\right)t - \left(\tilde{\gamma}_{\bar{c}} - \gamma_0\right)DT_t\left(\tilde{\tau}\right),$$

so for $s \in [0,1]$

$$T^{-1/2}\left(\tilde{u}_{\lfloor Ts \rfloor} - \hat{u}_{\lfloor Ts \rfloor}\right) = T^{-1/2}\left(\hat{\alpha}_{\bar{c}} - \tilde{\alpha}_{\bar{c}}\right) + T^{1/2}\left(\hat{\beta}_{\bar{c}} - \tilde{\beta}_{\bar{c}}\right)T^{-1}\lfloor Ts \rfloor$$

$$+ T^{1/2}\left(\hat{\gamma}_{\bar{c}} - \tilde{\gamma}_{\bar{c}}\right)T^{-1}DT_{\lfloor Ts \rfloor}\left(\tau_0\right)$$

$$- T^{1/2}\left(\tilde{\gamma}_{\bar{c}} - \gamma_0\right)T^{-1}\left(DT_{\lfloor Ts \rfloor}\left(\tilde{\tau}\right) - DT_{\lfloor Ts \rfloor}\left(\tau_0\right)\right).$$

27

Each of the first three of these terms are $O_p\left(T^{-1/2}\right)$ by (A.11) in Lemma 4. In the fourth term, we have (using $\tilde{\tau} < \tau_0$ without loss of generality)

$$DT_t\left(\tilde{\tau}\right) - DT_t\left(\tau_0\right) = \begin{cases} 0, & t \leq \lfloor \tilde{\tau}T \rfloor \\ t - \lfloor \tilde{\tau}T \rfloor, & \lfloor \tilde{\tau}T \rfloor < t \leq \lfloor \tau_0 T \rfloor \\ \lfloor \tau_0 T \rfloor - \lfloor \tilde{\tau}T \rfloor, & \lfloor \tau_0 T \rfloor < t \leq T \end{cases}$$

so

$$T^{-1}\left(DT_{\lfloor Ts \rfloor}\left(\tilde{\tau}\right) - DT_{\lfloor Ts \rfloor}\left(\tau_0\right)\right) = \begin{cases} 0, & \lfloor Ts \rfloor \leq \lfloor \tilde{\tau}T \rfloor \\ \left(\lfloor Ts \rfloor - \lfloor \tilde{\tau}T \rfloor\right)/T, & \lfloor \tilde{\tau}T \rfloor < \lfloor Ts \rfloor \leq \lfloor \tau_0 T \rfloor \\ \left(\lfloor \tau_0 T \rfloor - \lfloor \tilde{\tau}T \rfloor\right)/T, & \lfloor \tau_0 T \rfloor < \lfloor Ts \rfloor \leq T. \end{cases}$$

The component for $\lfloor \tau_0 T \rfloor < \lfloor Ts \rfloor \leq T$ is $O_p\left(T^{-1}\right)$ by (A.21) while the component for $\lfloor \tilde{\tau}T \rfloor < \lfloor Ts \rfloor \leq \lfloor \tau_0 T \rfloor$ satisfies

$$0 \leq \frac{\lfloor Ts \rfloor - \lfloor \tilde{\tau}T \rfloor}{T} \leq \frac{\lfloor \tau_0 T \rfloor - \lfloor \tilde{\tau}T \rfloor}{T} = O_p\left(T^{-1}\right).$$

Clearly these results all hold uniformly in s.

(ii) We first write

$$T^{-1} \sum_{t=p+1}^{T} \Delta\tilde{u}_{t-i}\Delta\tilde{u}_{t-j} - T^{-1} \sum_{t=p+1}^{T} \Delta\hat{u}_{t-i}\Delta\hat{u}_{t-j}$$

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

Substituting

$$\Delta\tilde{u}_t - \Delta\hat{u}_t = \left(\hat{\beta}_{\bar{c}} - \tilde{\beta}_{\bar{c}}\right) + \left(\hat{\gamma}_{\bar{c}} - \tilde{\gamma}_{\bar{c}}\right) DU_t\left(\tau_0\right) - \left(\tilde{\gamma}_{\bar{c}} - \gamma_0\right)\left(DU_t\left(\tilde{\tau}\right) - DU_t\left(\tau_0\right)\right), \quad \text{(A.22)}$$

into the first of these terms gives

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

$$= \left(\hat{\beta}_{\bar{c}} - \tilde{\beta}_{\bar{c}}\right) T^{-1} \sum_{t=p-i+1}^{T-i} \Delta\tilde{u}_{t-j+i} + \left(\hat{\gamma}_{\bar{c}} - \tilde{\gamma}_{\bar{c}}\right) T^{-1} \sum_{t=p-i+1}^{T-i} DU_t\left(\tau_0\right) \Delta\tilde{u}_{t-j+i}$$

$$\quad - \left(\tilde{\gamma}_{\bar{c}} - \gamma_0\right) T^{-1} \sum_{t=p-i+1}^{T-i} \left(DU_t\left(\tilde{\tau}\right) - DU_t\left(\tau_0\right)\right) \Delta\tilde{u}_{t-j+i}$$

$$= \left(\hat{\beta}_{\bar{c}} - \tilde{\beta}_{\bar{c}}\right) T^{-1} \left(\tilde{u}_{T-j} - \tilde{u}_{p-j-1}\right) + \left(\hat{\gamma}_{\bar{c}} - \tilde{\gamma}_{\bar{c}}\right) T^{-1} \left(\tilde{u}_{T-j} - \tilde{u}_{\lfloor \tau_0 T \rfloor}\right)$$

$$\quad - \left(\tilde{\gamma}_{\bar{c}} - \gamma_0\right) T^{-1} \left(\tilde{u}_{\lfloor \tau_0 T \rfloor} - \tilde{u}_{\lfloor \tilde{\tau}T \rfloor}\right)$$

28

which is obviously $O_p\left(T^{-1}\right)$, uniformly in i and $j \leq p$. The second term is similarly $O_p\left(T^{-1}\right)$.

(iii) We first write

$$T^{-1} \sum_{t=p+1}^{T} \tilde{u}_{t-1}\Delta\tilde{u}_{t-i} - T^{-1} \sum_{t=p+1}^{T} \hat{u}_{t-1}\Delta\hat{u}_{t-i} = T^{-1} \sum_{t=p}^{T-1} \left(\tilde{u}_t - \hat{u}_t\right)\Delta\tilde{u}_{t-i+1}$$

$$+ \; T^{-1} \sum_{t=p-i+1}^{T-i} \hat{u}_{t-1+i}\left(\Delta\tilde{u}_t - \Delta\hat{u}_t\right) \quad \text{(A.23)}$$

The first term can be written as

$$T^{-1} \sum_{t=p}^{T-1} \left(\tilde{u}_t - \hat{u}_t\right)\Delta\tilde{u}_{t-i+1} = T^{-1}\left(\tilde{u}_{T-1} - \hat{u}_{T-1}\right)\tilde{u}_{T-i} - T^{-1}\left(\tilde{u}_p - \hat{u}_p\right)\tilde{u}_{p-i}$$

$$-T^{-1} \sum_{t=p+1}^{T-1} \tilde{u}_{t-i}\left(\Delta\tilde{u}_t - \Delta\hat{u}_t\right)$$

the first two terms of which are $O_p\left(T^{-1/2}\right)$ by Lemma 5(i), while the third term is essentially the same as the second term in (A.23). Therefore we just need to show that the second term of (A.23) is of $O_p\left(T^{-1/2}\right)$. Substituting (A.22) into the second term of (A.23) gives

$$T^{-1} \sum_{t=p-i+1}^{T-i} \hat{u}_{t-1+i}\left(\Delta\tilde{u}_t - \Delta\hat{u}_t\right) = T^{1/2}\left(\hat{\beta}_{\bar{c}} - \tilde{\beta}_{\bar{c}}\right)T^{-3/2} \sum_{t=p-i+1}^{T-i} \hat{u}_{t-1+i}$$

$$+ T^{1/2}\left(\hat{\gamma}_{\bar{c}} - \tilde{\gamma}_{\bar{c}}\right)T^{-3/2} \sum_{t=\lfloor\tau_0 T\rfloor+1}^{T-i} \hat{u}_{t-1+i}$$

$$+ T^{1/2}\left(\tilde{\gamma}_{\bar{c}} - \gamma_0\right)T^{-3/2} \sum_{t=\lfloor\tilde{\tau}T\rfloor+1}^{\lfloor\tau_0 T\rfloor} \hat{u}_{t-1+i},$$

The first two terms of which are clearly of $O_p\left(T^{-1/2}\right)$ from (A.11). For the last term, we will show that

$$T^{-3/2} \sum_{t=\lfloor\tilde{\tau}T\rfloor+1}^{\lfloor\tau_0 T\rfloor} \hat{u}_{t-1+i} = O_p\left(T^{-1}\right). \quad \text{(A.24)}$$

For convenience we will just show (A.24) under the null with $c = 0$, though of course it will hold with $c > 0$ as well. We know that $\tilde{\tau}$ is $O_p\left(T^{-1}\right)$ consistent so that for any $\varepsilon > 0$ there exists a $B_\varepsilon > 0$ such that $\Pr\left(\left|\tilde{\tau} - \tau_0\right| > T^{-1}B_\varepsilon\right) < \varepsilon$ for all large enough T. Since we are only concerned with $\tilde{\tau} < \tau_0$ in this proof, we note that

$$\Pr\left(\left|T^{-3/2} \sum_{t=\lfloor\tilde{\tau}T\rfloor+1}^{\lfloor\tau_0 T\rfloor} \hat{u}_{t-1+i}\right| \geq \sup_{\tau\in[\tau_0-T^{-1}B_\varepsilon,\tau_0]} \left|T^{-3/2} \sum_{t=\lfloor\tau T\rfloor+1}^{\lfloor\tau_0 T\rfloor} \hat{u}_{t-1+i}\right|\right) \leq \varepsilon, \quad \text{(A.25)}$$

29

so we just need to derive the order of the right hand term in this inequality. Now

$$\hat{u}_t = u_t - X_t\left(\tau_0\right)'\left(\hat{\theta}_{\bar{c}} - \theta_0\right),$$

and $u_t = \sum_{j=1}^{t}\varepsilon_j$ under the null. Applying the BN decomposition to $\varepsilon_t = c\left(L\right)\eta_t$ gives $\varepsilon_t = c\left(1\right)\eta_t - \Delta\varepsilon_t^*$ and

$$T^{-3/2}\sum_{t=\lfloor\tau T\rfloor+1}^{\lfloor\tau_0 T\rfloor}u_{t-1} = c\left(1\right)T^{-3/2}\sum_{t=\lfloor\tau T\rfloor+1}^{\lfloor\tau_0 T\rfloor}\sum_{j=1}^{t-1}\eta_j + T^{-3/2}\sum_{t=\lfloor\tau T\rfloor+1}^{\lfloor\tau_0 T\rfloor}\left(\varepsilon_{t-1}^* - \varepsilon_1^*\right).$$

The second term is clearly $O_p\left(T^{-1}\right)$ while the first can be written

$$T^{-3/2}\sum_{t=\lfloor\tau T\rfloor+1}^{\lfloor\tau_0 T\rfloor}\sum_{j=1}^{t-1}\eta_j = T^{-3/2}\sum_{j=1}^{\lfloor\tau_0 T\rfloor-1}\sum_{t=(j+1)\vee(\lfloor\tau T\rfloor+1)}^{\lfloor\tau_0 T\rfloor}\eta_j = T^{-3/2}\sum_{j=1}^{\lfloor\tau_0 T\rfloor-1}\left(\lfloor\tau_0 T\rfloor - j\vee\lfloor\tau T\rfloor\right)\eta_j.$$

For any $\tau \in \left[\tau_0 - T^{-1}B_\varepsilon, \tau_0\right]$ this has mean zero and variance

$$T^{-3}\sum_{j=1}^{\lfloor\tau_0 T\rfloor-1}\left(\lfloor\tau_0 T\rfloor - (j\vee\lfloor\tau T\rfloor)\right)^2 \;\leq\; \frac{\lfloor\tau_0 T\rfloor}{T}\left(\frac{\lfloor\tau_0 T\rfloor - \lfloor\tau T\rfloor}{T}\right)^2$$

$$\leq\; \frac{\lfloor\tau_0 T\rfloor}{T}\left(\frac{\lfloor\tau_0 T\rfloor - \lfloor(\tau_0 - T^{-1}B_\varepsilon)T\rfloor}{T}\right)^2$$

$$= \frac{\lfloor\tau_0 T\rfloor}{T}\left(\frac{B_\varepsilon}{T} - \frac{\tau_0 T - \lfloor\tau_0 T\rfloor}{T} + \frac{(\tau_0 - T^{-1}B_\varepsilon)T - \lfloor(\tau_0 - T^{-1}B_\varepsilon)T\rfloor}{T}\right)^2 = O\left(T^{-2}\right).$$

Thus $T^{-3/2}\sum_{t=\lfloor\tau T\rfloor+1}^{\lfloor\tau_0 T\rfloor}u_{t-1} = O_p\left(T^{-1}\right)$ uniformly in $\tau \in \left[\tau_0 - T^{-1}B_\varepsilon, \tau_0\right]$. Next, since $D_T\left(\hat{\theta}_{\bar{c}} - \theta_0\right) = O_p\left(1\right)$ by (A.10), we just consider

$$T^{-3/2}\sum_{t=\lfloor\tau T\rfloor+1}^{\lfloor\tau_0 T\rfloor}D_T^{-1}X_t\left(\tau_0\right) = T^{-3/2}\sum_{t=\lfloor\tau T\rfloor+1}^{\lfloor\tau_0 T\rfloor}\begin{pmatrix}1\\ T^{-1/2}t\\ T^{-1/2}DT_t\left(\tau_0\right)\end{pmatrix}$$

$$= \begin{pmatrix}T^{-3/2}\left(\lfloor\tau_0 T\rfloor - \lfloor\tau T\rfloor\right)\\ \frac{1}{2}T^{-2}\left(\lfloor\tau_0 T\rfloor^2 + \lfloor\tau_0 T\rfloor - \lfloor\tau T\rfloor^2 - \lfloor\tau T\rfloor\right)\\ 0\end{pmatrix}.$$

The first element of the vector is $O\left(T^{-3/2}\right)$ for $\tau \in \left[\tau_0 - T^{-1}B_\varepsilon, \tau_0\right]$ while the second term satisfies

$$\frac{1}{2}T^{-2}\left(\lfloor\tau_0 T\rfloor^2 + \lfloor\tau_0 T\rfloor - \lfloor\tau T\rfloor^2 - \lfloor\tau T\rfloor\right)$$

$$= \frac{1}{2}\left(\frac{\lfloor\tau_0 T\rfloor - \lfloor\tau T\rfloor}{T}\right)\left(\frac{\lfloor\tau_0 T\rfloor + \lfloor\tau T\rfloor + 1}{T}\right) = O\left(T^{-1}\right).$$

30

We have therefore shown that $\sup_{\tau \in [\tau_0 - T^{-1}B_\varepsilon, \tau_0]} \left| T^{-3/2} \sum_{t=\lfloor \tau T \rfloor + 1}^{\lfloor \tau_0 T \rfloor} \hat{u}_{t-1+i} \right| = O_p(T^{-1})$ and, by (A.25), that (A.24) holds too.

Proof of Lemma 6

(i) As noted in Berk (1974, p.492), we can write

$$
\left\| \frac{\tilde{U}_p' \tilde{U}_p}{T} - \frac{\hat{U}_p' \hat{U}_p}{T} \right\| \leq \left(\sum_{i=1}^{p} \sum_{j=1}^{p} \left(T^{-1} \sum_{t=p+1}^{T} \Delta \tilde{u}_{t-i} \Delta \tilde{u}_{t-j} - T^{-1} \sum_{t=p+1}^{T} \Delta \tilde{u}_{t-i} \Delta \tilde{u}_{t-j} \right)^2 \right)^{1/2}
$$
$$
= O_p\left(pT^{-1} \right) \tag{A.26}
$$

from Lemma 5(ii). By standard manipulations (see, for example, Equation (2.15) of Berk, 1974), we can write

$$
\left(1 - \left\| \frac{\tilde{U}_p' \tilde{U}_p}{T} - \frac{\hat{U}_p' \hat{U}_p}{T} \right\| \left\| \left(\frac{\hat{U}_p' \hat{U}_p}{T} \right)^{-1} \right\| \right) \left(\left\| \left(\frac{\tilde{U}_p' \tilde{U}_p}{T} \right)^{-1} \right\| - \left\| \left(\frac{\hat{U}_p' \hat{U}_p}{T} \right)^{-1} \right\| \right)
$$
$$
\leq \left\| \left(\frac{\hat{U}_p' \hat{U}_p}{T} \right)^{-1} \right\| \left\| \frac{\tilde{U}_p' \tilde{U}_p}{T} - \frac{\hat{U}_p' \hat{U}_p}{T} \right\| \left\| \left(\frac{\hat{U}_p' \hat{U}_p}{T} \right)^{-1} \right\|.
$$

The right hand side of this inequality is $O_p\left(pT^{-1} \right)$ by (A.12) and (A.26). Similarly the first term on the left hand side is positive with probability approaching one. Thus

$$
\left\| \left(\frac{\tilde{U}_p' \tilde{U}_p}{T} \right)^{-1} \right\| - \left\| \left(\frac{\hat{U}_p' \hat{U}_p}{T} \right)^{-1} \right\| \xrightarrow{p} 0
$$

which establishes the result in part (i).

(ii) The stated result follows from (A.13) and the proof of part (vi), below.

(iii) First observe that

$$
\left\| \frac{\hat{U}_p' \Delta \hat{u}}{T} \right\| = \left(\sum_{j=1}^{p} \left(T^{-1} \sum_{t=p+1}^{T} \Delta \hat{u}_{t-j} \Delta \hat{u}_t \right)^2 \right)^{1/2} = O_p\left(p^{1/2} \right)
$$

since each of $T^{-1} \sum_{t=p+1}^{T} \Delta \hat{u}_{t-j} \Delta \hat{u}_t = O_p(1)$ for each j. Then the result that $\left\| T^{-1} \tilde{U}_p' \Delta \tilde{u} \right\| = O_p\left(p^{1/2} \right)$ follows from this and the proof of part (iv), below.

(iv) We write

$$
\left\| \frac{\tilde{U}_p' \Delta \tilde{u}}{T} - \frac{\hat{U}_p' \Delta \hat{u}}{T} \right\| = \left(\sum_{j=1}^{p} \left(T^{-1} \sum_{t=p+1}^{T} \Delta \tilde{u}_{t-j} \Delta \tilde{u}_t - T^{-1} \sum_{t=p+1}^{T} \Delta \hat{u}_{t-j} \Delta \hat{u}_t \right)^2 \right)^{1/2}
$$
$$
= O_p\left(p^{1/2} T^{-1} \right)
$$

31

from Lemma 5(ii).

(v) Simple reorganisation gives

$$
\left\| \left(\frac{\tilde{U}_p' \tilde{U}_p}{T} \right)^{-1} - \left(\frac{\hat{U}_p' \hat{U}_p}{T} \right)^{-1} \right\| \leq \left\| \left(\frac{\tilde{U}_p' \tilde{U}_p}{T} \right)^{-1} \right\| \left\| \frac{\tilde{U}_p' \tilde{U}_p}{T} - \frac{\hat{U}_p' \hat{U}_p}{T} \right\| \left\| \left(\frac{\hat{U}_p' \hat{U}_p}{T} \right)^{-1} \right\|
$$

$$
= O_p \left(p T^{-1} \right)
$$

from (A.26), (A.12) and part (i), above.

(vi) We write

$$
\left\| \frac{\tilde{u}_{-1}' \tilde{U}_p}{T} - \frac{\hat{u}_{-1}' \hat{U}_p}{T} \right\| = \left(\sum_{j=1}^{p} \left(T^{-1} \sum_{t=p+1}^{T} \Delta \tilde{u}_{t-j} \tilde{u}_{t-1} - T^{-1} \sum_{t=p+1}^{T} \Delta \hat{u}_{t-j} \hat{u}_{t-1} \right)^2 \right)^{1/2}
$$

which is $O_p \left(p^{1/2} T^{-1/2} \right)$ from Lemma 5(iii).

A.3 Proof of Lemma 2

The partial sums of the DGP in matrix form is written with an obvious notation as

$$
w = Z_{\tau_0} \theta_0 + s. \tag{A.27}
$$

We also partition Z_τ as $Z_\tau := (Z_1, z_{2,\tau})$ so that

$$
w = Z_1 \theta_{1,0} + z_{2,\tau} \gamma_0 + s,
$$

and the Wald test statistic for $\gamma_0 = 0$ can be written

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

where $\bar{P}_1 := I_T - Z_1 \left(Z_1' Z_1 \right)^{-1} Z_1'$ and $\bar{P}_\tau := I_T - Z_\tau \left(Z_\tau' Z_\tau \right)^{-1} Z_\tau'$.

We first derive the asymptotic theory for $W_T \left(\tau_0 \right)$. If $\tau = \tau_0$ we can use $w' \bar{P}_{\tau_0} w = s' \bar{P}_{\tau_0} s$ and $\bar{P}_{\tau_0} = \bar{P}_1 - \bar{P}_1 z_{2,\tau_0} \left(z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0} \right)^{-1} z_{2,\tau_0} \bar{P}_1$ to write

$$
W_T \left(\tau_0 \right) = \frac{\left(\gamma z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0} + z_{2,\tau_0}' \bar{P}_1 s \right)^2}{\left(z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0} \right) \left(s' \bar{P}_1 s \right) - \left(z_{2,\tau_0}' \bar{P}_1 s \right)^2}. \tag{A.28}
$$

In all cases we have the limit

$$
T^{-5} z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0} \rightarrow \int_0^1 Z_{2.1,\tau_0} \left(r \right)^2 dr,
$$

where $Z_{2.1,\tau_0}(r)$ is the residual process from a projection of $Z_{2,\tau_0}(r)$ on $Z_1(r)$.

(i) Suppose $\gamma_0 = 0$. In this case

$$T^{-9/2} z'_{2,\tau_0} \bar{P}_1 s \xrightarrow{d} \omega_\varepsilon \int_0^1 Z_{2.1,\tau_0}(r) S_1(r) \, dr,$$

and

$$T^{-4} s' \bar{P}_1 s \xrightarrow{d} \omega_\varepsilon^2 \int_0^1 S_1(r)^2 \, dr.$$

Substitution into (A.28) gives

$$\begin{aligned}
W_T(\tau_0) &= \frac{\left(T^{-9/2} z'_{2,\tau_0} \bar{P}_1 s\right)^2}{\left(T^{-5} z'_{2,\tau_0} \bar{P}_1 z_{2,\tau_0}\right)\left(T^{-4} s' \bar{P}_1 s\right) - \left(T^{-9/2} z'_{2,\tau_0} \bar{P}_1 s\right)^2} \\
&\xrightarrow{d} \frac{\left(\int_0^1 Z_{2.1,\tau_0}(r) S_1(r) \, dr\right)^2}{\int_0^1 Z_{2.1,\tau_0}(r)^2 \, dr \int_0^1 S_1(r)^2 \, dr - \left(\int_0^1 Z_{2.1,\tau_0}(r) S_1(r) \, dr\right)^2} \\
&= \frac{\int_0^1 S_1(r)^2 \, dr}{\int_0^1 S_{1,\tau_0}(r)^2 \, dr} - 1.
\end{aligned}$$

(ii) If $\gamma_0 \neq 0$ then

$$\begin{aligned}
T^{-1} W_T(\tau_0) &= \frac{\left(\gamma_0 T^{-5} z'_{2,\tau_0} \bar{P}_1 z_{2,\tau_0} + T^{-5} z'_{2,\tau_0} \bar{P}_1 s\right)^2}{\left(T^{-5} z'_{2,\tau_0} \bar{P}_1 z_{2,\tau_0}\right)\left(T^{-4} s' \bar{P}_1 s\right) - \left(T^{-9/2} z'_{2,\tau_0} \bar{P}_1 s\right)^2} \qquad \text{(A.29)} \\
&\xrightarrow{d} \frac{\gamma_0^2 \int_0^1 Z_{2.1,\tau_0}(r)^2 \, dr}{\int_0^1 Z_{2.1,\tau_0}(r)^2 \, dr \int_0^1 S_1(r)^2 \, dr - \left(\int_0^1 Z_{2.1,\tau_0}(r) S_1(r) \, dr\right)^2} \\
&= \gamma_0^2 \left(\int S_{1,\tau_0}(r)^2 \, dr\right)^{-1}.
\end{aligned}$$

(iii) For $\gamma_0 = 0$, the statistic with appropriate normalisation is

$$W_T(\tilde{\tau}) = \frac{\left(T^{-9/2} z'_{2,\tilde{\tau}} \bar{P}_1 s\right)^2}{\left(T^{-5} z'_{2,\tilde{\tau}} \bar{P}_1 z_{2,\tilde{\tau}}\right)\left(T^{-4} s' \bar{P}_1 s\right) - \left(T^{-9/2} z'_{2,\tilde{\tau}} \bar{P}_1 s\right)^2}.$$

Using the fact that $\tilde{\tau} = O_p(1)$ (see, Nunes $et\ al.$, 1995, p.741), we will show that $T^{-9/2} z'_{2,\tilde{\tau}} \bar{P}_1 s = O_p(1)$ and $T^{-5} z'_{2,\tilde{\tau}} \bar{P}_1 z_{2,\tilde{\tau}} = O_p(1)$ while $T^{-4} s' \bar{P}_1 s \xrightarrow{d} \omega_1^2 \int_0^1 S_1(r)^2 \, dr$ as before.

Now,

$$T^{-9/2} z'_{2,\tilde{\tau}} \bar{P}_1 s = \frac{z'_{2,\tilde{\tau}} s}{T^{9/2}} - \frac{z'_{2,\tilde{\tau}} Z_1 D_{1,T}^{-1}}{T^3} \left(\frac{D_{1,T}^{-1} Z_1' Z_1 D_{1,T}^{-1}}{T}\right)^{-1} \frac{D_{1,T}^{-1} Z_1' s}{T^{5/2}},$$

where $D_{1,T} = \mathsf{diag}\,(T, T^2)$. Here,

$$
\begin{aligned}
\left| \frac{z'_{2,\tilde\tau} s}{T^{9/2}} \right| &= \left| T^{-9/2} \sum_{t=1}^{T} \left(\sum_{j=1}^{t} 0 \vee (j - \tilde\tau) \right) s_t \right| \\
&\le T^{-2} \sum_{j=0}^{T} j \cdot T^{-5/2} \sum_{t=1}^{T} |s_t| = O_p\,(1)\,.
\end{aligned}
$$

Also,

$$
\begin{aligned}
\frac{D_{1,T}^{-1} Z'_1 z_{2,\tilde\tau}}{T^3} &= T^{-3} \sum_{t=1}^{T} \begin{pmatrix} t/T \\ \tfrac{1}{2} t^2/T^2 \end{pmatrix} \left(\sum_{j=1}^{t} 0 \vee (j - \tilde\tau) \right) \\
&\le T^{-3} \sum_{t=\tilde\tau}^{T} \sum_{j=\tilde\tau}^{t} j \le T^{-3} \sum_{t=1}^{T} \sum_{j=1}^{t} j = O\,(1)\,,
\end{aligned}
$$

while $T^{-1} D_{1,T}^{-1} Z'_1 Z_1 D_{1,T}^{-1}$ and $T^{-5/2} D_{1,T}^{-1} Z'_1 s$ do not involve $\tilde\tau$ and are clearly $O_p\,(1)$ as well. Thus $T^{-9/2} z'_{2,\tilde\tau} \bar{P}_1 s = O_p\,(1)$. Similarly

$$
T^{-5} z'_{2,\tilde\tau} \bar{P}_1 z_{2,\tilde\tau} = \frac{z'_{2,\tilde\tau} z_{2,\tilde\tau}}{T^5} - \frac{z'_{2,\tilde\tau} Z_1 D_{1,T}^{-1}}{T^3} \left(\frac{D_{1,T}^{-1} Z'_1 Z_1 D_{1,T}^{-1}}{T} \right)^{-1} \frac{D_{1,T}^{-1} Z'_1 z_{2,\tilde\tau}}{T^3},
$$

where

$$
\begin{aligned}
\frac{z'_{2,\tilde\tau} z_{2,\tilde\tau}}{T^5} &= T^{-5} \sum_{t=1}^{T} \left(\sum_{j=1}^{t} 0 \vee (j - \tilde\tau) \right)^2 \\
&= T^{-5} \sum_{t=0}^{T-\tilde\tau} \left(\sum_{j=0}^{t} j \right)^2 \\
&\le T^{-5} \sum_{t=0}^{T} \left(\sum_{j=0}^{t} j \right)^2 = O_p\,(1)\,.
\end{aligned}
$$

Thus $T^{-5} z'_{2,\tilde\tau} \bar{P}_1 z_{2,\tilde\tau} = O_p\,(1)$ and hence $W_T\,(\tilde\tau) = O_p\,(1)\,.$

(iv) Now, $\tilde\tau - \tau_0 = O_p\,(T^{-1})$ and in view of (A.29) we will show that (for $\tilde\tau \ge \tau_0$ is sufficient): (a) $T^{-5} z'_{2,\tau_0} \bar{P}_1 z_{2,\tau_0} - T^{-5} z'_{2,\tilde\tau} \bar{P}_1 z_{2,\tilde\tau} = o_p\,(1)$; (b) $T^{-9/2} z'_{2,\tau_0} \bar{P}_1 s - T^{-9/2} z'_{2,\tilde\tau} \bar{P}_1 s = o_p\,(1)$, while obviously $T^{-5} z'_{2,\tilde\tau} \bar{P}_1 s = o_p\,(1)$. For (a), we will show

$$
\frac{D_{1,T}^{-1} Z'_1 z_{2,\tau_0}}{T^3} - \frac{D_{1,T}^{-1} Z'_1 z_{2,\tilde\tau}}{T^3} = o_p\,(1) \tag{A.30}
$$

and

$$
\frac{z'_{2,\tau_0} z_{2,\tau_0}}{T^5} - \frac{z'_{2,\tilde\tau} z_{2,\tilde\tau}}{T^5} = \frac{z'_{2,\tau_0} (z_{2,\tau_0} - z_{2,\tilde\tau})}{T^5} + \frac{(z_{2,\tau_0} - z_{2,\tilde\tau})' z_{2,\tilde\tau}}{T^5} = o_p\,(1)\,.
$$

34

First

$$\left| \frac{D_{1,T}^{-1} Z_1' z_{2,\tau_0}}{T^3} - \frac{D_{1,T}^{-1} Z_1' z_{2,\tilde{\tau}}}{T^3} \right| = \left| T^{-3} \sum_{t=1}^{T} \begin{pmatrix} t/T \\ \frac{1}{2} t^2/T^2 \end{pmatrix} \sum_{j=1}^{t} (DT_j(\tau_0) - DT_j(\tilde{\tau})) \right|$$

$$\leq T^{-2} \sum_{t=1}^{T} |DT_t(\tau_0) - DT_t(\tilde{\tau})| \cdot T^{-1} \sum_{t=1}^{T} \begin{pmatrix} t/T \\ \frac{1}{2} t^2/T^2 \end{pmatrix}$$

$$= \left(T^{-1} \sum_{\tau_0 T < t \leq \tilde{\tau} T} \left(\frac{t}{T} - \tau_0 \right) + T^{-1} \sum_{\tilde{\tau} T < t \leq T} (\tilde{\tau} - \tau_0) \right) \cdot T^{-1} \sum_{t=1}^{T} \begin{pmatrix} t/T \\ \frac{1}{2} t^2/T^2 \end{pmatrix}$$

$$\leq (\tilde{\tau} - \tau_0) \cdot T^{-1} \sum_{t=1}^{T} \begin{pmatrix} t/T \\ \frac{1}{2} t^2/T^2 \end{pmatrix} = O_p(T^{-1}),$$

and second

$$\frac{z_{2,\tau_0}'(z_{2,\tau_0} - z_{2,\tilde{\tau}})}{T^5} = T^{-5} \sum_{t=1}^{T} \left(\sum_{i=1}^{t} DT_i(\tau_0) \right) \left(\sum_{j=1}^{t} (DT_j(\tau_0) - DT_j(\tilde{\tau})) \right)$$

$$\leq T^{-2} \sum_{j=1}^{T} |DT_j(\tau_0) - DT_j(\tilde{\tau})| \cdot T^{-3} \sum_{t=1}^{T} \left(\sum_{i=1}^{t} DT_i(\tau_0) \right)$$

$$\leq (\tilde{\tau} - \tau_0) \cdot T^{-3} \sum_{t=1}^{T} \left(\sum_{i=1}^{t} i \right) = O_p(T^{-1}),$$

and similarly for $T^{-5}(z_{2,\tau_0} - z_{2,\tilde{\tau}})' z_{2,\tilde{\tau}}$. Therefore $T^{-5} z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0} - T^{-5} z_{2,\tilde{\tau}}' \bar{P}_1 z_{2,\tilde{\tau}} = o_p(1)$. For (b), since we have already established (A.30), it will be sufficient to show that

$$\frac{z_{2,\tau_0}' s}{T^{9/2}} - \frac{z_{2,\tilde{\tau}}' s}{T^{9/2}} = o_p(1),$$

which follows from

$$\left| \frac{z_{2,\tau_0}' s}{T^{9/2}} - \frac{z_{2,\tilde{\tau}}' s}{T^{9/2}} \right| = \left| T^{-9/2} \sum_{t=1}^{T} \sum_{j=1}^{t} (DT_j(\tau_0) - DT_j(\tilde{\tau})) s_t \right|$$

$$\leq T^{-2} \sum_{j=1}^{T} |DT_j(\tau_0) - DT_j(\tilde{\tau})| \cdot T^{-5/2} \sum_{t=1}^{T} |s_t|$$

$$\leq (\tilde{\tau} - \tau_0) \cdot T^{-5/2} \sum_{t=1}^{T} |s_t| = O_p(T^{-1}).$$

Together, (a) and (b) imply that $T^{-1} W_T(\tau_0) - T^{-1} W_T(\tilde{\tau}) = o_p(1)$, as required.

35

A.4 Proof of Lemma 3

(i) When $\gamma_0 = 0$, $\tilde{\tau} = O_p(1)$, from Nunes *et al.* (1995, p.741). Consequently, $\bar{\tau} = (1 - \bar{\lambda})O_p(1) = O_p(T^{-1/2})$, since $1 - \bar{\lambda} = 1 - \exp(-gT^{-1/2}W_T(\tilde{\tau})) = T^{-1/2}gW_T(\tilde{\tau}) - \frac{g^2}{2}T^{-1}W_T(\tilde{\tau})^2 + \cdots = O_p(T^{-1/2})$, because $W_T(\tilde{\tau}) = O_p(1)$.

(ii) When $\gamma_0 \neq 0$, we find that $T(\bar{\tau} - \tau_0) = T(\tilde{\tau} - \tau_0) - T\bar{\lambda}\tilde{\tau} = O_p(1) + o_p(1)O_p(1) = O_p(1)$, since $T\bar{\lambda} = T\exp(-gT^{-1/2}W_T(\tilde{\tau})) = T(1 - T^{-1/2}gW_T(\tilde{\tau}) + \frac{g^2}{2}T^{-1}W_T(\tilde{\tau})^2 - \cdots)^{-1} = o_p(1)$, as $T^{-1}W_T(\tilde{\tau})^2 = O_p(T)$.

References

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.

Beveridge, S. and Nelson, C.R. (1981). A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. Journal of Monetary Economics 7, 151–174

Carrion-i-Silvestre, J.L. Kim, D. and Perron, P. (2007). GLS-based unit root tests with multiple structural breaks both under the null and the alternative hypotheses. Downloadable from `http://people.bu.edu/perron/`

Cavaliere, G. and Georgiev, I. (2007). Testing for unit roots in autoregressions with multiple level shifts. Forthcoming in Econometric Theory.

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

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

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

Hajek, J. and Renyi, A. (1955). A generalization of an inequality of Kolomogorov, Acta. Math. Acad. Sci. Hung. 6, 281-284.

Hansen, B.E. (2007a) Least squares model averaging. Econometrica 75, 1175-1189.

Hansen, B.E. (2007b) Averaging estimators for regressions with a possible structural break. Downloadable from `http://www.nottingham.ac.uk/economics/granger centre/conference/program.htm`

Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2007a). Unit Root Testing in Practice: Dealing with Uncertainty over the Trend and Initial Condition. Granger Centre for Time Series Econometrics, Working Paper 07/03.

Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2007b). Simple, robust and powerful tests of the breaking trend hypothesis. Forthcoming in Econometric Theory.

Kim, D. and Perron, P. (2006). Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. Boston University Economics Working Paper 2006/52.

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

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

Phillips, P.C.B. (1991a). To criticize the critics: an objective Bayesian analysis of stochastic trends. Journal of Applied Econometrics 6, 333-64,

Phillips, P.C.B. (1991b). Bayesian routes and unit roots: de rebus prioribus semper est disputandum. Journal of Applied Econometrics 6, 435-73.

Phillips, P.C.B. and Ploberger, W. (1994). Posterior odds testing for a unit root with data-based model selection. Econometric Theory 10, 774-808.

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

Perron, P. and Yabu, T. (2006). Testing for shifts in trend with an integrated or stationary noise component. Unpublished Manuscript, Department of Economics, Boston University.

Perron, P. and Yabu, T. (2007). Estimating Deterministic Trends with an Integrated or Stationary Noise Component. Boston University Economics Working Paper 2007/25.

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

Sayginsoy, O. and Vogelsang, T.J. (2004). Powerful tests of structural change that are robust to strong serial correlation. Working Paper 04/08, Department of Economics, SUNY at Albany.

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. Critical Values and QD De-trending Parameters

| | 0.10 level | | | | 0.05 level | | | | 0.01 level | | | |
| | | Critical values | | | | Critical values | | | | Critical values | | |
τ_0	c_0	$T=150$	$T=300$	$T=\infty$	c_0	$T=150$	$T=300$	$T=\infty$	c_0	$T=150$	$T=300$	$T=\infty$
0.15	−13.4	−3.13	−3.11	−3.09	−17.6	−3.42	−3.40	−3.37	−26.2	−4.01	−3.95	−3.93
0.20	−13.8	−3.17	−3.15	−3.12	−17.8	−3.46	−3.44	−3.40	−26.6	−4.04	−3.99	−3.95
0.25	−14.0	−3.21	−3.18	−3.15	−18.2	−3.50	−3.46	−3.42	−26.6	−4.09	−4.02	−3.96
0.30	−14.2	−3.24	−3.19	−3.16	−18.4	−3.53	−3.48	−3.43	−26.8	−4.10	−4.04	−3.98
0.35	−14.4	−3.26	−3.21	−3.16	−18.6	−3.54	−3.49	−3.43	−27.0	−4.12	−4.04	−3.99
0.40	−14.4	−3.28	−3.22	−3.16	−18.4	−3.55	−3.50	−3.44	−27.0	−4.12	−4.04	−3.98
0.45	−14.4	−3.28	−3.21	−3.15	−18.4	−3.56	−3.50	−3.44	−26.6	−4.11	−4.04	−3.99
0.50	−14.2	−3.28	−3.21	−3.14	−18.2	−3.55	−3.49	−3.42	−26.8	−4.12	−4.05	−3.96
0.55	−14.0	−3.26	−3.20	−3.13	−18.0	−3.54	−3.49	−3.41	−26.6	−4.12	−4.04	−3.96
0.60	−13.8	−3.24	−3.18	−3.11	−17.6	−3.52	−3.47	−3.39	−26.0	−4.10	−4.03	−3.93
0.65	−13.4	−3.22	−3.16	−3.08	−17.4	−3.50	−3.44	−3.37	−25.8	−4.08	−4.01	−3.91
0.70	−13.2	−3.19	−3.13	−3.04	−17.0	−3.47	−3.41	−3.34	−25.4	−4.05	−3.98	−3.87
0.75	−12.6	−3.15	−3.09	−3.00	−16.6	−3.44	−3.37	−3.29	−25.0	−4.01	−3.94	−3.83
0.80	−12.2	−3.10	−3.03	−2.96	−16.0	−3.39	−3.32	−3.24	−24.4	−3.96	−3.89	−3.79
0.85	−11.6	−3.02	−2.96	−2.89	−15.2	−3.32	−3.26	−3.17	−23.6	−3.87	−3.83	−3.74
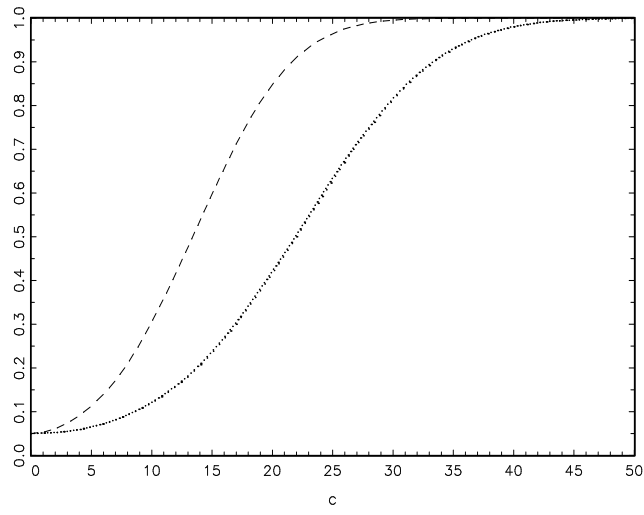
Figure 1. Asymptotic size and local power: $\gamma_0 = 0$; $t(\bar{\tau})$, $t_P(\tilde{\tau})$: $--$, $t_C(\hat{\tau})$: ⋯⋯ , t_{ZA}: $\cdot\ \cdot\ \cdot$

F.1
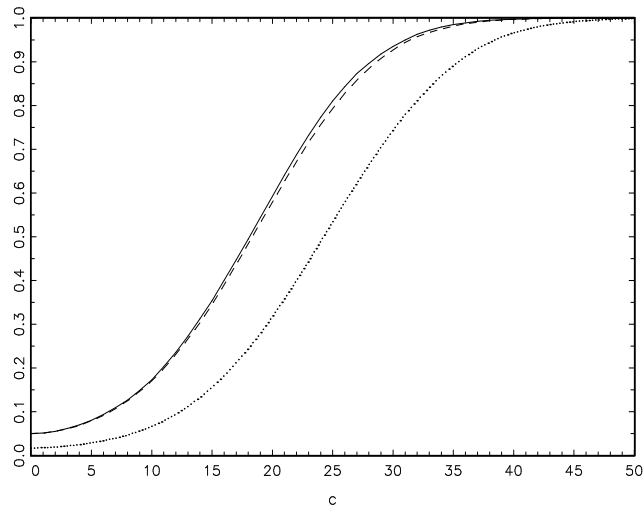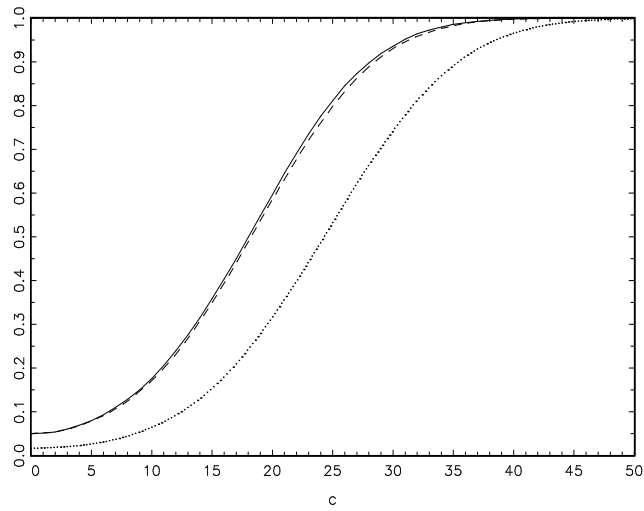
(a) $\tau_0 = 0.3$

(b) $\tau_0 = 0.5$
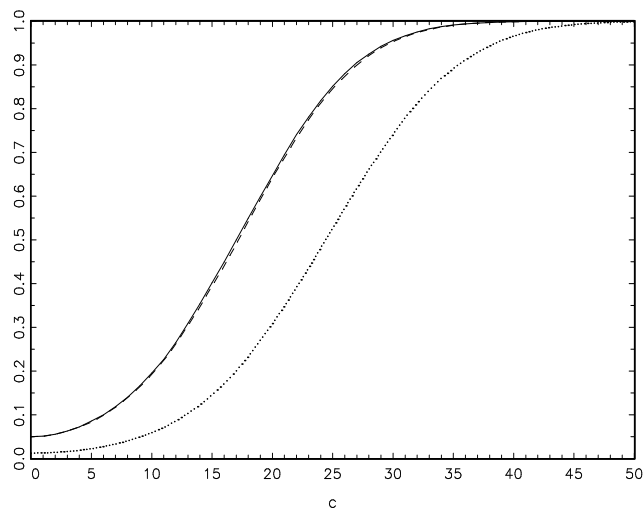
(c) $\tau_0 = 0.7$

Figure 2. Asymptotic size and local power: $\gamma_0 \neq 0$;
Gaussian power envelope: ——— , $t(\bar{\tau})$, $t_P(\tilde{\tau})$: $--$, $t_C(\hat{\tau})$: ·····

F.2
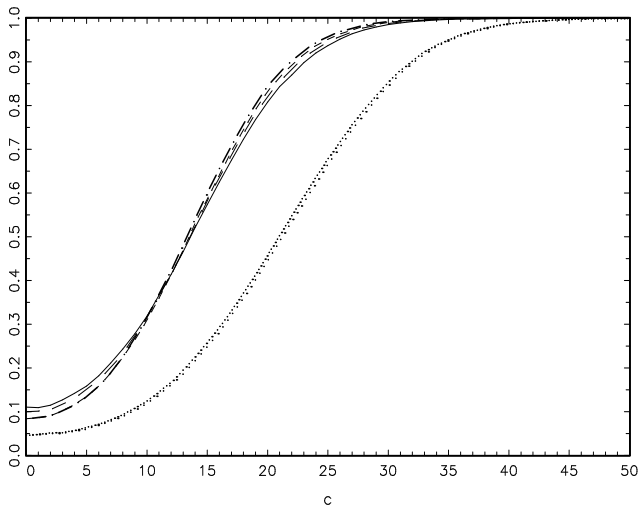
(a) $T = 150$
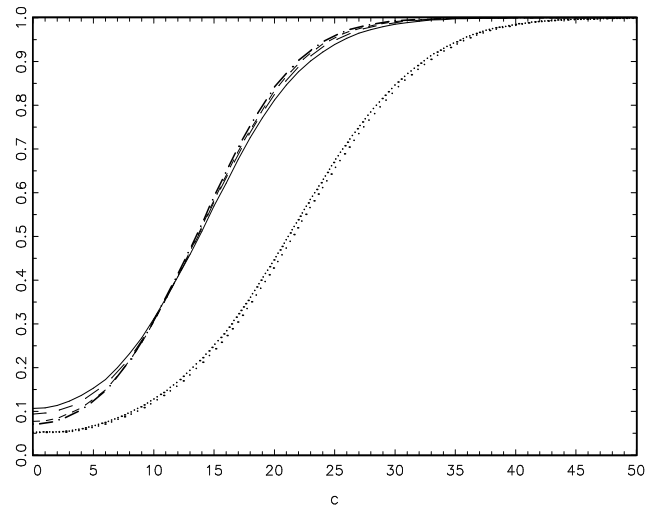
(b) $T = 300$

Figure 3. Finite sample size and power: $\gamma_0 = 0$;
$t(\bar{\tau}, 1.5)$: - - - , $t(\bar{\tau}, 3)$: $--$, $t(\bar{\tau}, 6)$: —— , $t_P(\tilde{\tau})$: $-\cdot-$, $t_C(\hat{\tau})$: ⋯⋯ , t_{ZA}: · · ·
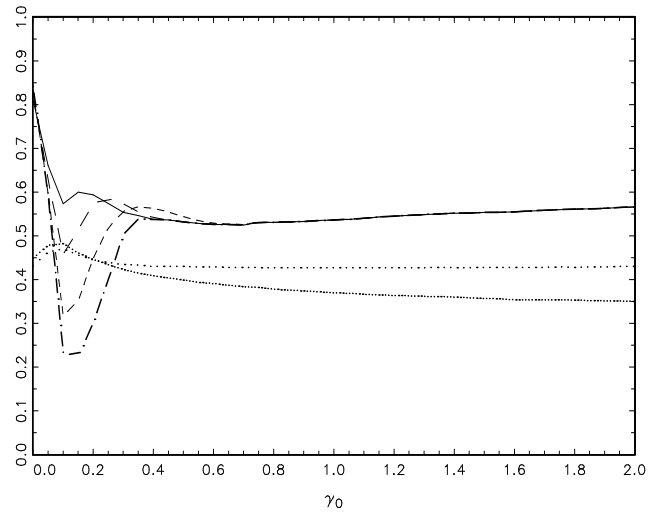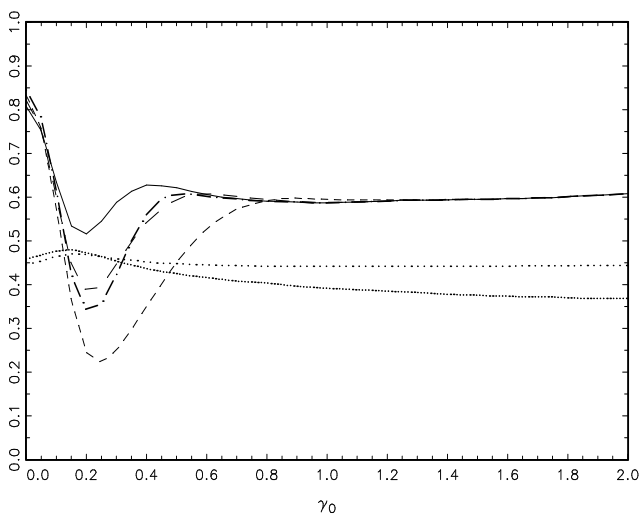
F.3
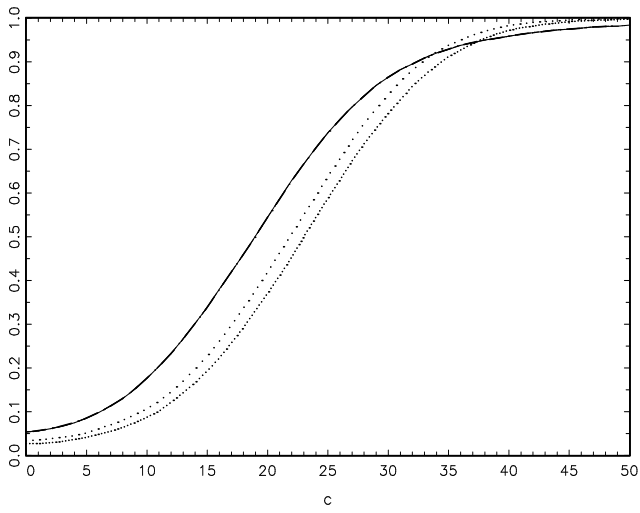
(a) $\tau_0 = 0.3$, $T = 150$
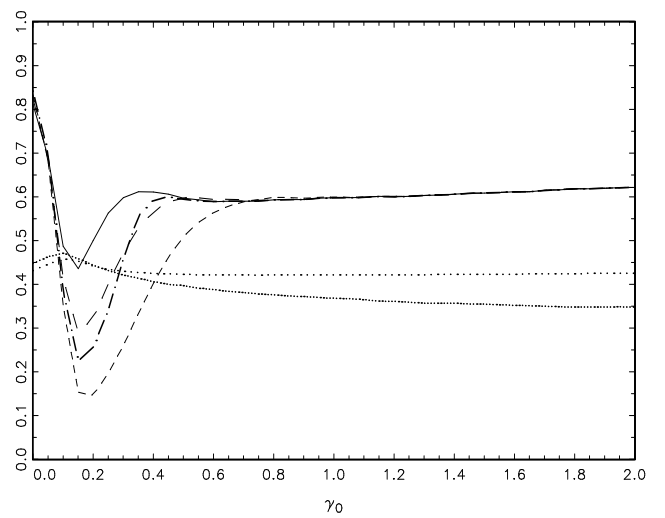
(b) $\tau_0 = 0.3$, $T = 300$

(c) $\tau_0 = 0.5$, $T = 150$

(d) $\tau_0 = 0.5$, $T = 300$

(c) $\tau_0 = 0.7$, $T = 150$

(d) $\tau_0 = 0.7$, $T = 300$

Figure 4. Finite sample size and power: $\gamma_0 = 1$;
$t(\bar{\tau}, 1.5)$: - - - , $t(\bar{\tau}, 3)$: $--$, $t(\bar{\tau}, 6)$: —— , $t_P(\tilde{\tau})$: $- \cdot -$, $t_C(\hat{\tau})$: ⋯⋯ , t_{ZA}: · · ·

F.4
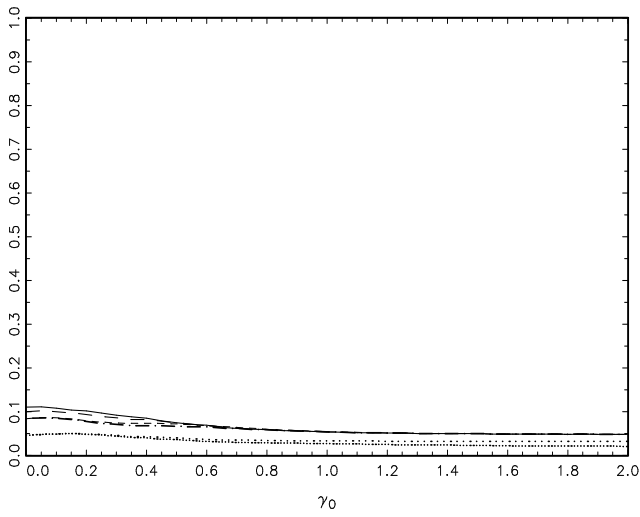
(a) $\tau_0 = 0.3$, $T = 150$

(b) $\tau_0 = 0.3$, $T = 300$

(c) $\tau_0 = 0.5$, $T = 150$

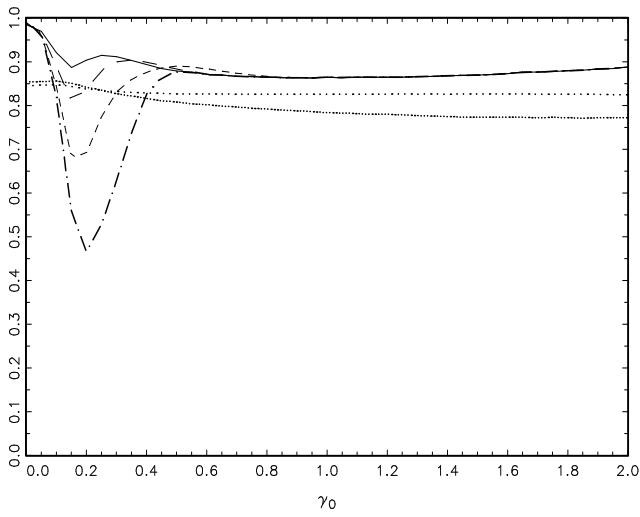(d) $\tau_0 = 0.5$, $T = 300$
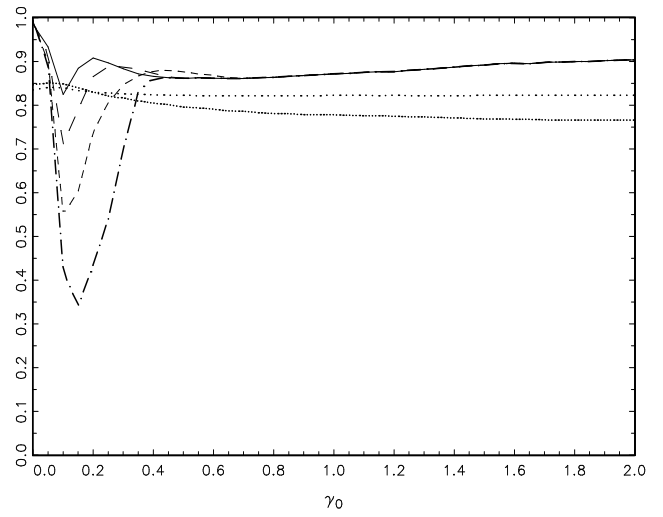
(c) $\tau_0 = 0.7$, $T = 150$

(d) $\tau_0 = 0.7$, $T = 300$

Figure 5. Finite sample size: $c = 0$;

$t(\bar{\tau}, 1.5)$: - - - , $t(\bar{\tau}, 3)$: $--$, $t(\bar{\tau}, 6)$: —— , $t_P(\tilde{\tau})$: $-\cdot-$, $t_C(\hat{\tau})$: ····· , t_{ZA}: · · ·

F.5
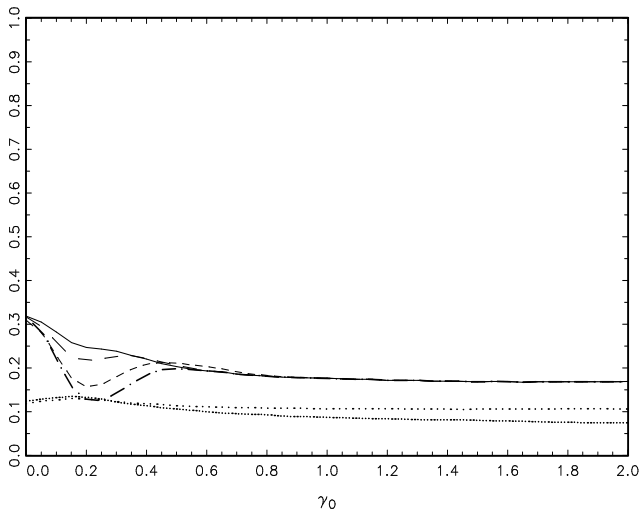
(a) $\tau_0 = 0.3$, $T = 150$

(b) $\tau_0 = 0.3$, $T = 300$

(c) $\tau_0 = 0.5$, $T = 150$

(d) $\tau_0 = 0.5$, $T = 300$

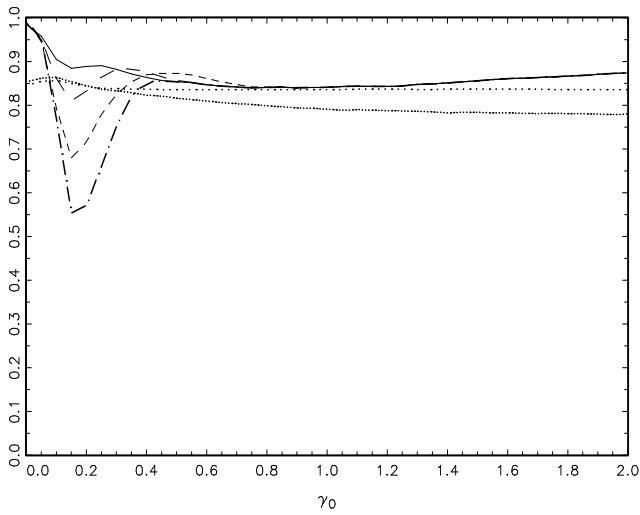(c) $\tau_0 = 0.7$, $T = 150$

(d) $\tau_0 = 0.7$, $T = 300$

Figure 6. Finite sample power: $c = 10$;
$t(\bar{\tau}, 1.5)$: - - - , $t(\bar{\tau}, 3)$: − − , $t(\bar{\tau}, 6)$: ——— , $t_P(\tilde{\tau})$: − · − , $t_C(\hat{\tau})$: ····· , t_{ZA}: · · ·
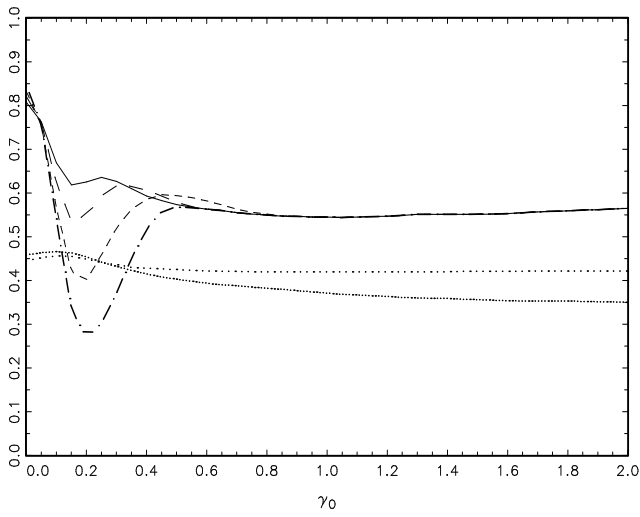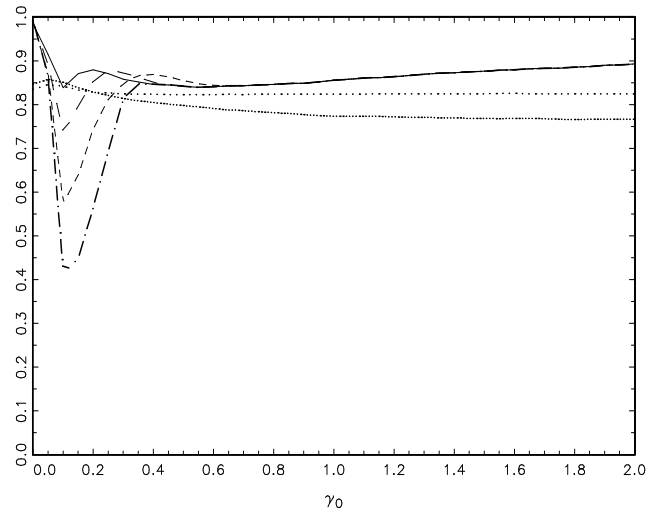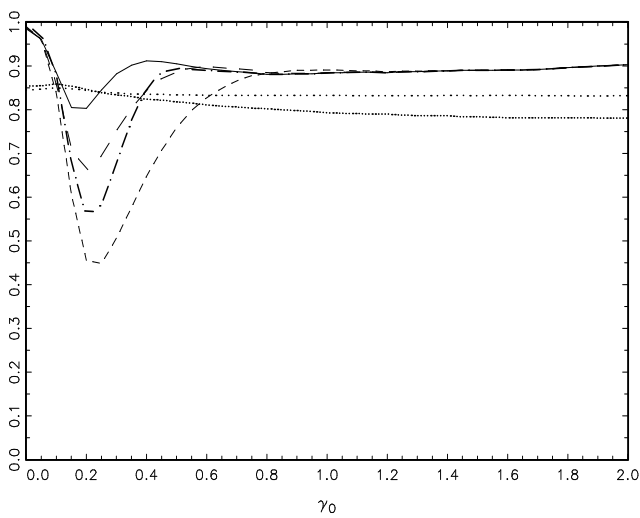
(a) $\tau_0 = 0.3$, $T = 150$
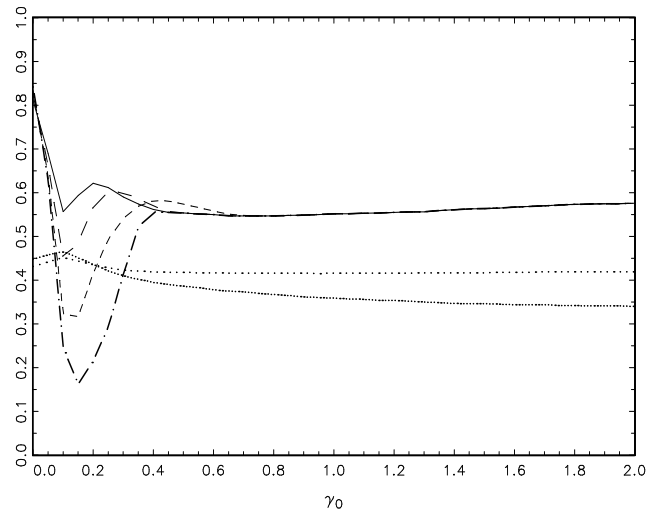
(b) $\tau_0 = 0.3$, $T = 300$
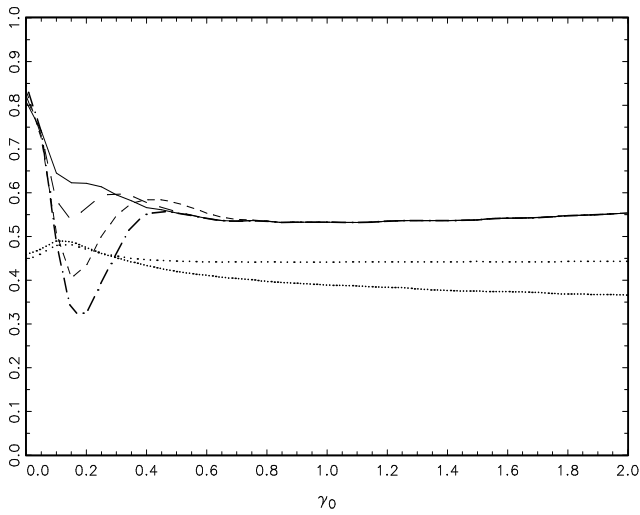
(c) $\tau_0 = 0.5$, $T = 150$

(d) $\tau_0 = 0.5$, $T = 300$

(c) $\tau_0 = 0.7$, $T = 150$

(d) $\tau_0 = 0.7$, $T = 300$

Figure 7. Finite sample power: $c = 20$;
$t(\bar{\tau}, 1.5)$: - - - , $t(\bar{\tau}, 3)$: $--$, $t(\bar{\tau}, 6)$: —— , $t_P(\tilde{\tau})$: $- \cdot -$, $t_C(\hat{\tau})$: ····· , t_{ZA}: · · ·

F.7

(a) $\tau_0 = 0.3$, $T = 150$
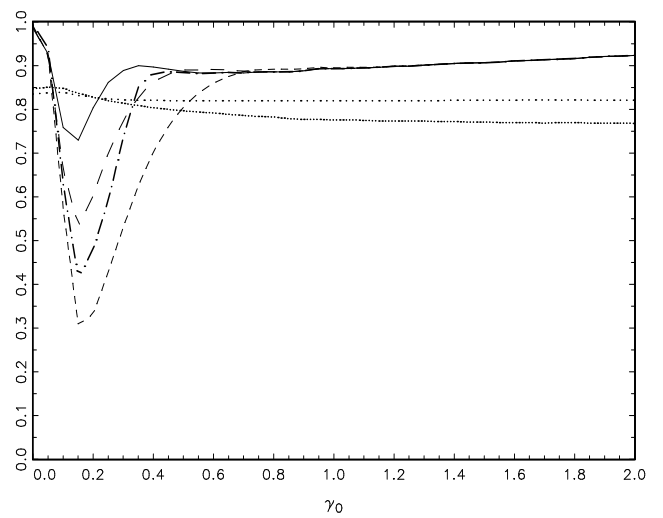
(b) $\tau_0 = 0.3$, $T = 300$

(c) $\tau_0 = 0.5$, $T = 150$

(d) $\tau_0 = 0.5$, $T = 300$

(c) $\tau_0 = 0.7$, $T = 150$

(d) $\tau_0 = 0.7$, $T = 300$

Figure 8. Finite sample power: $c = 30$;
$t(\bar{\tau}, 1.5)$: - - - , $t(\bar{\tau}, 3)$: − − , $t(\bar{\tau}, 6)$: —— , $t_P(\tilde{\tau})$: − · − , $t_C(\hat{\tau})$: ····· , t_{ZA}: · · ·

F.8