

On Infimum Dickey-Fuller Unit Root Tests Allowing for a Trend Break Under The Null*

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

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

October 2012

Abstract

Trend breaks appear to be prevalent in macroeconomic time series. Consequently, to avoid the catastrophic impact that unmodelled trend breaks have on power it is standard empirical practice to employ unit root tests which allow for such effects. A popularly applied approach is the infimum ADF-type test. Its appeal has endured with practitioners despite results which show that the infimum ADF statistic diverges to $-\infty$ as the sample size diverges, with the consequence that the test has an asymptotic size of unity when a break in trend is present under the unit root null hypothesis. The result for additive outlier-type breaks in trend (but not intercept) is refined and shows that divergence to $-\infty$ occurs only when the true break fraction is smaller than $2/3$. An alternative testing strategy based on the maximum of the original infimum statistic and the corresponding statistic constructed using the time-reversed sample data is considered.

Keywords: Unit root test; trend break; Minimum Dickey-Fuller test.

*Address correspondence to: David Harvey, School of Economics, University of Nottingham, Nottingham NG7 2RD, UK. Tel.: +44 (0)115 9515481. Fax: +44 (0)115 9514159. E-mail: dave.harvey@nottingham.ac.uk

1 Introduction

Macroeconomic series appear to often be characterized by broken trend functions; see, *inter alia*, Stock and Watson (1996,1999,2005) and Perron and Zhu (2005). In a seminal paper, Perron (1989) shows that failure to account for trend breaks present in the data results in unit root tests with zero power, even asymptotically. Consequently, when testing for a unit root it has become a matter of regular practice to allow for this kind of deterministic structural change. While Perron (1989) initially treated the location of the break as known, subsequent approaches have focused on the case where the break location is unknown and is chosen via a data-dependent method; see, *inter alia*, Zivot and Andrews (1992) [ZA], Banerjee *et al.* (1992) and Perron (1997); see also Pitarakis (2012).

Of these endogenised approaches, the testing methodology proposed by ZA has been widely used by practitioners (for a recent example, see Alexeev and Maynard, 2012). The approach suggested by ZA is to calculate the ADF t -ratio-type statistic of Perron (1989) for all candidate break points within a trimmed range and to then form a test based on the infimum (most negative) of this sequence of statistics. This infimum test is simple to compute and, by selecting the statistic within the sequence which gives most weight to the alternative, follows an established approach to such problems in econometrics.

A significant drawback with the infimum approach, however, is that it is predicated on the maintained hypothesis that no break in trend occurs under the unit root null hypothesis. This assumption needs to be made in order for the infimum statistic to have a pivotal limiting null distribution. Investigating what happens when this maintained assumption does not hold, results presented in Vogelsang and Perron (1998) show that, for both sudden additive outlier (AO) breaks and slowly evolving innovational outlier (IO) breaks, when a trend break of fixed non-zero magnitude occurs under the unit root null, so the infimum statistics diverge to $-\infty$ as the sample size diverges and, hence, cause the tests to have asymptotic size of unity.

In this paper we revisit this issue, focusing on AO-type breaks in the trend (but not intercept). Our primary contribution is to refine the theoretical results given in Vogelsang and Perron (1998), showing that the divergence of the infimum statistic to $-\infty$ occurs only when the true break fraction, τ_0 say, is smaller than $2/3$. We also briefly consider an alternative testing strategy based on the maximum of the original infimum statistic and the corresponding statistic constructed using the time-reversed sample data. We find that such an approach appears to offer considerable improvements in finite sample size control relative to the original test, while retaining attractive power properties.

In the following ‘ $[\cdot]$ ’ denotes the integer part of its argument, ‘ \Rightarrow ’ and ‘ \xrightarrow{P} ’ denote weak convergence and convergence in probability, respectively, in each case as the sample size diverges to $+\infty$, ‘ $x := y$ ’ (‘ $x =: y$ ’) indicates that x is defined by y (y is defined by x), and ‘ $1(\cdot)$ ’ denotes the indicator function.

2 The Model and the Infimum Unit Root Test

We consider a univariate time series $\{y_t\}$ generated by the AO DGP,

$$y_t = \mu + \beta t + \gamma DT_t(\tau_0) + u_t, \quad t = 1, \dots, T, \quad (1)$$

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad t = 2, \dots, T \quad (2)$$

where, for a generic fraction τ , $DT_t(\tau) := 1(t > \lfloor \tau T \rfloor)(t - \lfloor \tau T \rfloor)$ in (1), and τ_0 is an (unknown) putative trend break fraction, with associated break magnitude γ . The break fraction is assumed to be such that $\tau_0 \in \Lambda$, where $\Lambda := [\tau_L, \tau_U]$ with $0 < \tau_L < \tau_U < 1$; the fractions τ_L and τ_U representing trimming parameters. In (2), $\{u_t\}$ is an unobserved mean zero stochastic process, initialised such that $u_1 = o_p(T^{1/2})$. Also, for simplicity of exposition, we will assume that ε_t in (2) is an independent and identically distributed sequence with mean zero, variance σ_ε^2 and finite fourth moment. The theoretical results given in the paper would continue to hold under a more general weak dependence assumption provided the Dickey-Fuller-type unit root regression in (3) below was augmented with k lags of the dependent variable and where k satisfies the usual condition that $1/k + k^3/T \rightarrow 0$ as $T \rightarrow \infty$.

We examine the problem of testing the unit root null hypothesis $H_0 : \rho = 1$, against the alternative, $H_1 : \rho < 1$, without assuming knowledge of where, or indeed if, the trend break occurs in the DGP. Let \hat{u}_t denote the residual from fitting the OLS regression of y_t on $z_t := [1, t, DT_t(\tau)]'$ (we suppress the dependence of \hat{u}_t and any associated OLS estimators on τ for convenience of notation), i.e.

$$\hat{u}_t := y_t - \hat{\mu} - \hat{\beta}t - \hat{\gamma}DT_t(\tau)$$

and let $t_{\hat{\phi}}(\tau)$ denote the Dickey-Fuller t -ratio for testing $\phi = 0$ in the fitted OLS regression

$$\Delta \hat{u}_t = \hat{\phi} \hat{u}_{t-1} + \hat{\varepsilon}_t \quad (3)$$

that is

$$t_{\hat{\phi}}(\tau) := \frac{\sum_{t=2}^T \Delta \hat{u}_t \hat{u}_{t-1}}{\sqrt{\hat{\sigma}_\varepsilon^2 \sum_{t=2}^T \hat{u}_{t-1}^2}}$$

with $\hat{\sigma}_\varepsilon^2 := (T-2)^{-1} \sum_{t=2}^T \hat{\varepsilon}_t^2$. Then the infimum ZA-type procedure we consider is based on the statistic

$$ZA_{AO} := \inf_{\tau \in \Lambda} t_{\hat{\phi}}(\tau).$$

3 Limiting Behaviour of ZA_{AO}

In order to derive the large sample behaviour of the ZA_{AO} statistic we must first evaluate the limiting behaviour of the $t_{\hat{\phi}}(\tau)$ statistic for $\tau \neq \tau_0$. This is provided in Theorem 1.

Theorem 1 *Let y_t be generated according to (1) and (2), with $\gamma = \kappa \sigma_\varepsilon$, $\kappa \neq 0$. Then for $\tau \neq \tau_0$*

$$T^{-1/2} t_{\hat{\phi}}(\tau) \xrightarrow{p} \frac{\kappa^2 \{N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2\}}{2\sqrt{\kappa^2 \{1 + \kappa^2 M(\tau_0, \tau)\} \int_0^1 N(r, \tau_0, \tau)^2 dr}}$$

with

$$\begin{aligned}
N(r, \tau_0, \tau) &:= \mathbb{I}_{\tau_0}^r(r - \tau_0) - P_1 - P_2 r - \mathbb{I}_{\tau}^r P_3(r - \tau), \\
M(\tau_0, \tau) &:= L(\tau_0, \tau) - \frac{\{N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2\}^2}{2 \int_0^1 N(r, \tau_0, \tau)^2 dr}, \\
L(\tau_0, \tau) &:= (1 - \tau_0) + P_2^2 + (1 - \tau)P_3^2 - 2(1 - \tau_0)P_2 - 2(1 - \tau_0 - \mathbb{I}_{\tau_0}^{\tau}(\tau - \tau_0))P_3 + 2(1 - \tau)P_2P_3
\end{aligned}$$

and

$$\begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} := \begin{bmatrix} 1 & \frac{1}{2} & \frac{(1-\tau)^2}{2} \\ \frac{1}{2} & \frac{1}{3} & \frac{(1-\tau)^2(2+\tau)}{6} \\ \frac{(1-\tau)^2}{2} & \frac{(1-\tau)^2(2+\tau)}{6} & \frac{(1-\tau)^3}{3} \end{bmatrix}^{-1} \begin{bmatrix} \frac{(1-\tau_0)^2}{2} \\ \frac{(1-\tau_0)^2(2+\tau_0)}{6} \\ \frac{\{1-\tau_0-\mathbb{I}_{\tau_0}^{\tau}(\tau-\tau_0)\}^2\{2+\tau_0-3\tau+4\mathbb{I}_{\tau_0}^{\tau}(\tau-\tau_0)\}}{6} \end{bmatrix}$$

where $\mathbb{I}_x^y := 1(y > x)$.

Notice that when $\tau = \tau_0$, we find that $[P_1, P_2, P_3] = [0, 0, 1]$ and $N(r, \tau_0, \tau_0) = 0$; consequently, the limit of $T^{-1/2}t_{\hat{\phi}}(\tau_0)$ is undefined. We know, however, from Kim and Perron (2009) that $t_{\hat{\phi}}(\tau)$ has a well-defined limit distribution under H_0 when τ is within an $o(T^{-1/2})$ neighbourhood of τ_0 , thus $t_{\hat{\phi}}(\tau) = O_p(1)$ here, and we will therefore consider the limit of $T^{-1/2}t_{\hat{\phi}}(\tau_0)$ to be zero. For $\tau \neq \tau_0$, Theorem 1 implies that $t_{\hat{\phi}}(\tau) = O_p(T^{1/2})$. Now since ZA_{AO} takes the minimum value of $t_{\hat{\phi}}(\tau)$ across all $\tau \in \Lambda$, the pertinent issue is the *sign* of the limit of $T^{-1/2}t_{\hat{\phi}}(\tau)$. More specifically, if, given a break at time τ_0 , $T^{-1/2}t_{\hat{\phi}}(\tau)$ is positive for *all* $\tau \in \Lambda$, then, since $t_{\hat{\phi}}(\tau) \xrightarrow{p} +\infty$ for all $\tau \neq \tau_0$, ZA_{AO} would not be minimised over this problem region, but rather for a value of τ within a shrinking neighbourhood of τ_0 . On the other hand, if $T^{-1/2}t_{\hat{\phi}}(\tau)$ is negative for *any* $\tau \in \Lambda$, then $t_{\hat{\phi}}(\tau) \xrightarrow{p} -\infty$ for some $\tau \neq \tau_0$, and consequently $ZA_{AO} \xrightarrow{p} -\infty$ also, resulting in unit asymptotic size. We now therefore examine the sign of the limit of $T^{-1/2}t_{\hat{\phi}}(\tau)$ as a function of τ_0 and τ .

The sign of the limit of $T^{-1/2}t_{\hat{\phi}}(\tau)$ is determined by the sign of $N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2$ as the other terms in the limit are unambiguously positive (as is clear from the Proof of Theorem 1, $1 + \kappa^2 M(\tau_0, \tau)$ is the limit of $\hat{\sigma}_{\varepsilon}^2/\sigma_{\varepsilon}^2$ and is therefore positive). Next note that we can write

$$\begin{aligned}
N(0, \tau_0, \tau) &= -P_1, \\
N(1, \tau_0, \tau) &= (1 - \tau_0) - P_1 - P_2 - P_3(1 - \tau)
\end{aligned}$$

and so

$$N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2 = \{(1 - \tau_0) - P_1 - P_2 - P_3(1 - \tau)\}^2 - P_1^2.$$

First consider the case where $\tau < \tau_0$. Here, we find (upon simplification)

$$\begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = \begin{bmatrix} \frac{(1-\tau_0)^2(\tau_0-\tau)}{2(1-\tau)} \\ -\frac{3(1-\tau_0)^2(\tau_0-\tau)}{2\tau(1-\tau)} \\ \frac{(1-\tau_0)^2(3\tau_0-\tau-2\tau\tau_0)}{2\tau(1-\tau)^3} \end{bmatrix}$$

yielding

$$N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2 = j_{\tau, \tau_0} h_{\tau, \tau_0} \tag{4}$$

where

$$\begin{aligned} j_{\tau, \tau_0} &:= (1 - \tau_0)^2 (\tau_0 - \tau)^2 (1 - \tau + \tau_0 - \tau) / 4(1 - \tau)^4, \\ h_{\tau, \tau_0} &:= 2\tau_0(1 - \tau) - (1 - \tau_0) \end{aligned}$$

with j_{τ, τ_0} always positive when $\tau < \tau_0$. Now, the function h_{τ, τ_0} is, for a given τ_0 , monotonically decreasing in τ , and since $\tau_L \leq \tau < \tau_0$ it is bounded by $[2\tau_0(1 - \tau_L) - (1 - \tau_0), (2\tau_0 - 1)(1 - \tau_0)]$. We then find that

$$\begin{aligned} \text{For } \tau_0 < 1/3 & \quad h_{\tau, \tau_0} < 0 \text{ for all } \tau_L \leq \tau < \tau_0, \\ \text{For } 1/3 \leq \tau_0 < 1/2 & \quad h_{\tau, \tau_0} \begin{cases} < 0 & \text{for } \tau_L \leq \tau < \frac{3\tau_0 - 1}{2\tau_0} \\ \geq 0 & \text{for } \frac{3\tau_0 - 1}{2\tau_0} \leq \tau < \tau_0 \end{cases}, \\ \text{For } \tau_0 \geq 1/2 & \quad h_{\tau, \tau_0} > 0 \text{ for all } \tau_L \leq \tau < \tau_0. \end{aligned}$$

Next, when $\tau > \tau_0$ we have

$$\begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = \begin{bmatrix} \frac{\tau_0(\tau - \tau_0)(-2\tau + \tau_0 + \tau\tau_0)}{2\tau^2} \\ \frac{(\tau - \tau_0)(2\tau^2 - \tau_0^2 + 2\tau\tau_0 - 3\tau\tau_0^2)}{2\tau^3} \\ \frac{\tau_0^2(-3\tau + \tau_0 + 2\tau\tau_0)}{2\tau^3(\tau - 1)} \end{bmatrix}$$

giving

$$N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2 = k_{\tau, \tau_0} l_{\tau, \tau_0} \quad (5)$$

where

$$\begin{aligned} k_{\tau, \tau_0} &:= \tau_0^2 (\tau - \tau_0)^2 (2\tau - \tau_0) / 4\tau^4, \\ l_{\tau, \tau_0} &:= \tau_0 - 2\tau(1 - \tau_0) \end{aligned}$$

with k_{τ, τ_0} always positive when $\tau > \tau_0$. The function l_{τ, τ_0} is, for a given τ_0 , monotonically decreasing in τ and since $\tau_0 < \tau \leq \tau_U$ it is bounded by $(\tau_0(2\tau_0 - 1), \tau_0 - 2\tau_U(1 - \tau_0))$. Then,

$$\begin{aligned} \text{For } \tau_0 < 1/2 & \quad l_{\tau, \tau_0} < 0 \text{ for all } \tau_0 < \tau \leq \tau_U, \\ \text{For } 1/2 \leq \tau_0 < 2/3 & \quad l_{\tau, \tau_0} \begin{cases} < 0 & \text{for } \tau_0 \leq \tau < \frac{\tau_0}{2(1 - \tau_0)} \\ \geq 0 & \text{for } \frac{\tau_0}{2(1 - \tau_0)} \leq \tau \leq \tau_U \end{cases}, \\ \text{For } \tau_0 \geq 2/3 & \quad l_{\tau, \tau_0} > 0 \text{ for all } \tau_0 < \tau \leq \tau_U. \end{aligned}$$

Drawing on the above results for (4) when $\tau < \tau_0$ and (5) when $\tau > \tau_0$, we can then write

$$\begin{aligned} \text{For } \tau_0 < 1/3 & \quad p \lim(T^{-1/2} t_{\hat{\phi}}(\tau)) < 0 \text{ for all } \tau_L \leq \tau \leq \tau_U, \\ \text{For } 1/3 \leq \tau_0 < 2/3 & \quad p \lim(T^{-1/2} t_{\hat{\phi}}(\tau)) \begin{cases} < 0 & \text{for some } \tau_L \leq \tau \leq \tau_U \\ \geq 0 & \text{for some } \tau_L \leq \tau \leq \tau_U \end{cases}, \\ \text{For } \tau_0 \geq 2/3 & \quad p \lim(T^{-1/2} t_{\hat{\phi}}(\tau)) > 0 \text{ for all } \tau_L \leq \tau \leq \tau_U. \end{aligned}$$

To further illustrate the behaviour of $T^{-1/2}t_{\hat{\phi}}(\tau)$, Figure 1 displays the regions in (τ, τ_0) space where $p \lim(T^{-1/2}t_{\hat{\phi}}(\tau))$ is positive/negative. Translating this behaviour into that of ZA_{AO} , which locates the minimum of $t_{\hat{\phi}}(\tau)$ across all $\tau \in \Lambda$, we consequently obtain that

$$ZA_{AO} \Rightarrow \begin{cases} -\infty & \tau_0 < 2/3 \\ O_p(1) & \tau_0 \geq 2/3. \end{cases}$$

We therefore find that, under H_0 , ZA_{AO} will spuriously reject with probability approaching one in the limit provided the true break fraction τ_0 lies below $2/3$. It will, however, not spuriously reject with probability one in the limit if τ_0 is $2/3$ or above. It is this second finding that refines the result presented in Vogelsang and Perron (1998), since the finding that spurious rejections of the null occur with probability one in the limit is here shown to depend on where the true break is located. (Note that, in contrast, the innovational outlier version of the statistic diverges to $-\infty$ for *all* τ_0 , as stated by Vogelsang and Perron (1998).) Of course, in an empirical setting where uncertainty exists as to the presence or location of a break, one is unlikely to have any confidence that the putative break will lie in the region $\tau_0 \in [2/3, \tau_U]$, and so the fundamental problem raised by Vogelsang and Perron (1998) of potentially serious over-sizing persists in the practical application of ZA_{AO} .

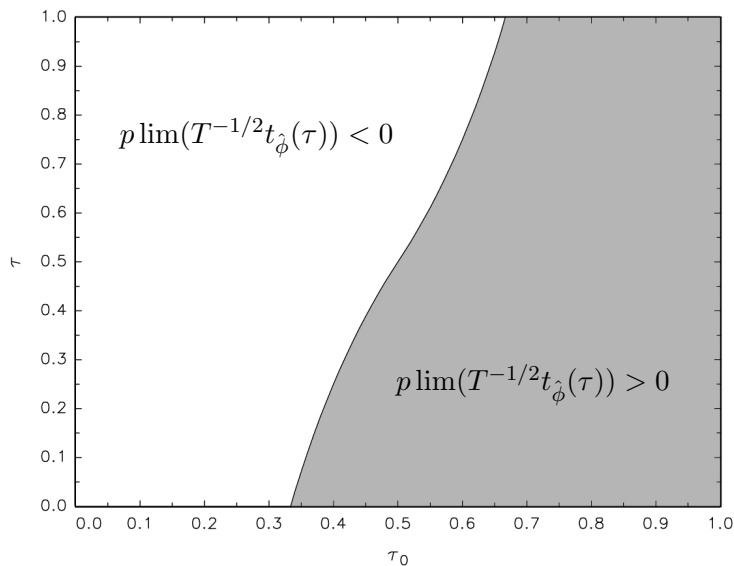


Figure 1. Sign of $p \lim(T^{-1/2}t_{\hat{\phi}}(\tau))$

4 Use of Time-Reversed Data

In order to consider how the asymptotic over-size problem region of a ZA_{AO} -type test might be reduced, we now consider the following improvisation. If we time-reverse the data, i.e. consider $\{y_{T-t+1}\}_{t=1}^T$ in place of $\{y_t\}_{t=1}^T$, then any break appearing in the first half of the original sample $\{y_t\}$ is translated into one occurring in the second half of $\{y_{T-t+1}\}$. Thus application of ZA_{AO} to $\{y_{T-t+1}\}$, which we denote by ZA'_{AO} , will deliver a test which does not spuriously reject in the limit with probability one when a break occurs in the first third of $\{y_t\}$. Of course, since in practice we have no information regarding

the location of a break, ZA'_{AO} does not achieve robustness, since here a break in the last two-thirds of $\{y_t\}$ would cause spurious rejection of the null by ZA'_{AO} . However, if we consider the *maximum* of ZA_{AO} and ZA'_{AO} (cf. Leybourne, 1995, in the context of unit root testing without allowance for a break in trend); that is,

$$ZA_{AO}^{\max} := \max(ZA_{AO}, ZA'_{AO})$$

the ZA_{AO} problem of spuriously rejecting the null with probability approaching one when $\tau_0 \in [\tau_L, 2/3]$ is for ZA_{AO}^{\max} restricted to the region $\tau_0 \in (1/3, 2/3)$. Under H_0 and for the case $\gamma = 0$, it is straightforward to show that

$$ZA_{AO}^{\max} \Rightarrow \max \left(\inf_{\tau \in \Lambda} \mathcal{Z}(\tau), \inf_{\tau \in \Lambda} \mathcal{Z}'(\tau) \right) \quad (6)$$

where

$$\mathcal{Z}(\tau) := \frac{K(1, \tau)^2 - K(0, \tau)^2 - 1}{2\sqrt{\int_0^1 K(r, \tau)^2 dr}}, \quad \mathcal{Z}'(\tau) := \frac{K'(1, \tau)^2 - K'(0, \tau)^2 - 1}{2\sqrt{\int_0^1 K'(r, \tau)^2 dr}}$$

with $K(r, \tau)$ and $K'(r, \tau)$ the continuous time residual processes from the projections of $W(r)$, and $W(1-r)$, respectively, onto the space spanned by $\{1, r, (r-\tau)\mathbb{I}_\tau^r\}$, where $W(r)$ is the Brownian motion process defined by $T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} \varepsilon_t \Rightarrow \sigma_\varepsilon W(r)$. Table 1 reports asymptotic nominal 0.10, 0.05 and 0.01 level critical values for ZA_{AO}^{\max} for a selection of commonly used trimming parameters. The critical values were obtained by direct simulation of (6), approximating the Wiener processes in the limiting functionals using $NIID(0, 1)$ random variates, with the integrals approximated by normalized sums of 2,000 steps. Unreported simulations, available from the authors on request, suggest that (i) ZA_{AO}^{\max} displays good finite sample size control for all τ_0 values, unless a large magnitude break occurs at $\tau_0 \in (1/3, 2/3)$ in an extremely large sample, and (ii) ZA_{AO}^{\max} has attractive power properties under H_1 , regardless of whether or not a break is actually present. Given that substantial over-size is only seen to occur for series that are unrepresentative of those encountered in typical economic applications, a pragmatic case could be made for using ZA_{AO}^{\max} , although it is difficult to fully justify such an approach, given that the test still has *asymptotic* size approaching one when $\tau_0 \in (1/3, 2/3)$.

Table 1. Asymptotic ξ -level critical values for the ZA_{AO}^{\max} test

$[\tau_L, \tau_U]$	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$
[0.05, 0.95]	-3.81	-4.07	-4.64
[0.10, 0.90]	-3.80	-4.06	-4.62
[0.15, 0.85]	-3.78	-4.04	-4.61
[0.20, 0.80]	-3.75	-4.02	-4.60

Acknowledgment

Financial support provided by the Economic and Social Research Council of the United Kingdom under research grant RES-000-22-3882 is gratefully acknowledged by the authors.

References

- Alexeev, V. and Maynard, A. (2012). Localized level crossing random walk test robust to the presence of structural breaks. *Computational Statistics and Data Analysis* 56, 3322-3344.
- 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 Economic Statistics* 10, 271-288.
- Kim, D. and Perron, P. (2009). Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. *Journal of Econometrics* 148, 1-13.
- Leybourne, S.J. (1995). Testing for unit roots using forward and reverse Dickey-Fuller regressions. *Oxford Bulletin of Economics and Statistics* 57, 559-571.
- 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 Zhu, X. (2005). Structural breaks with deterministic and stochastic trends. *Journal of Econometrics* 129, 65-119.
- Pitarakis, J.-Y. (2012). A joint test for structural stability and a unit root in autoregressions. *Computational Statistics and Data Analysis*, forthcoming.
- Stock, J.H. and Watson, M.W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics* 14, 11-30.
- Stock, J.H. and Watson, M.W. (1999). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. Engle, R.F. and White, H. (eds.), *Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive W.J. Granger*, Oxford: Oxford University Press, 1-44.
- Stock, J. and Watson, M.W. (2005). Implications of dynamic factor analysis for VAR models, NBER Working Paper #11467.
- Vogelsang, T.J. 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.

Appendix: Proof of Theorem 1

In what follows we can set $\mu = \beta = 0$ without loss of generality. When $\tau \neq \tau_0$ we have

$$\begin{aligned} T^{-2} \sum_{t=1}^T y_t &= T^{-2} \sum_{t=1}^T u_t + \kappa\sigma_\varepsilon T^{-2} \sum_{t=\tau_0 T+1}^T (t - \tau_0 T) \xrightarrow{p} \kappa\sigma_\varepsilon (1 - \tau_0)^2 / 2, \\ T^{-3} \sum_{t=1}^T t y_t &= T^{-3} \sum_{t=1}^T t u_t + \kappa\sigma_\varepsilon T^{-3} \sum_{t=\tau_0 T+1}^T t(t - \tau_0 T) \xrightarrow{p} \kappa\sigma_\varepsilon (1 - \tau_0)^2 (2 + \tau_0) / 6, \\ T^{-3} \sum_{t=\tau T+1}^T (t - \tau T) y_t &= T^{-3} \sum_{t=\tau T+1}^T (t - \tau T) u_t + \kappa\sigma_\varepsilon T^{-3} \sum_{t=\tau T+1}^T (t - \tau T) D T_t(\tau_0) \\ &\xrightarrow{p} \kappa\sigma_\varepsilon \{[1 - \tau_0 - \mathbb{I}_{\tau_0}^r(\tau - \tau_0)]^2 \{2 + \tau_0 - 3\tau + 4\mathbb{I}_{\tau_0}^r(\tau - \tau_0)\}\} / 6. \end{aligned}$$

Examining the limit behaviour of the residuals \hat{u}_t we obtain

$$\begin{aligned} \begin{bmatrix} T^{-1} \hat{\mu} \\ \hat{\beta} \\ \hat{\gamma} \end{bmatrix} &= \begin{bmatrix} 1 & T^{-2} \sum_{t=1}^T t & T^{-2} \sum_{t=\tau T+1}^T (t - \tau T) \\ T^{-2} \sum_{t=1}^T t & T^{-3} \sum_{t=1}^T t^2 & T^{-3} \sum_{t=\tau T+1}^T t(t - \tau T) \\ T^{-2} \sum_{t=\tau T+1}^T (t - \tau T) & T^{-3} \sum_{t=\tau T+1}^T t(t - \tau T) & T^{-3} \sum_{t=\tau T+1}^T (t - \tau T)^2 \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} T^{-2} \sum_{t=1}^T y_t \\ T^{-3} \sum_{t=1}^T t y_t \\ T^{-3} \sum_{t=\tau T+1}^T (t - \tau T) y_t \end{bmatrix} \\ &\xrightarrow{p} \begin{bmatrix} 1 & \frac{1}{2} & \frac{(1-\tau)^2}{2} \\ \frac{1}{2} & \frac{1}{3} & \frac{(1-\tau)^2(2+\tau)}{6} \\ \frac{(1-\tau)^2}{2} & \frac{(1-\tau)^2(2+\tau)}{6} & \frac{(1-\tau)^3}{3} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\kappa\sigma_\varepsilon(1-\tau_0)^2}{2} \\ \frac{\kappa\sigma_\varepsilon(1-\tau_0)^2(2+\tau_0)}{6} \\ \frac{\kappa\sigma_\varepsilon\{1-\tau_0-\mathbb{I}_{\tau_0}^r(\tau-\tau_0)\}^2\{2+\tau_0-3\tau+4\mathbb{I}_{\tau_0}^r(\tau-\tau_0)\}}{6} \end{bmatrix} \\ &= : \kappa\sigma_\varepsilon \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} \end{aligned}$$

giving

$$\begin{aligned} T^{-1} \hat{u}_{[rT]} &= T^{-1} y_{[rT]} - T^{-1} \hat{\mu} - T^{-1} \hat{\beta} [rT] - T^{-1} \hat{\gamma} \mathbb{I}_r^r([rT] - [\tau T]) \\ &= T^{-1} u_{[rT]} + \kappa\sigma_\varepsilon \mathbb{I}_{\tau_0}^r(r - \tau_0) - T^{-1} \hat{\mu} - \hat{\beta} r - \hat{\gamma} \mathbb{I}_r^r(r - \tau) \\ &\xrightarrow{p} \kappa\sigma_\varepsilon \{ \mathbb{I}_{\tau_0}^r(r - \tau_0) - P_1 - P_2 r - \mathbb{I}_r^r P_3 (r - \tau) \} =: \kappa\sigma_\varepsilon N(r, \tau_0, \tau). \end{aligned}$$

We also require

$$\begin{aligned} \Delta \hat{u}_t &= \Delta y_t - \hat{\beta} - \hat{\gamma} D U_t(\tau) \\ &= \kappa\sigma_\varepsilon D U_t(\tau_0) + \Delta u_t - \hat{\beta} - \hat{\gamma} D U_t(\tau) \\ \sum_{t=2}^T \Delta \hat{u}_t^2 &= \kappa^2 \sigma_\varepsilon^2 (T - \tau_0 T) + \sum_{t=2}^T \Delta u_t^2 + (T - 1) \hat{\beta}^2 + \hat{\gamma}^2 (T - \tau T) \\ &\quad + 2\kappa\sigma_\varepsilon \sum_{t=\tau_0 T+1}^T \Delta u_t - 2\kappa\sigma_\varepsilon (T - \tau_0 T) \hat{\beta} - 2\kappa\sigma_\varepsilon \hat{\gamma} \sum_{t=2}^T D U_t(\tau_0) D U_t(\tau) \\ &\quad - 2\hat{\beta} \sum_{t=2}^T \Delta u_t - 2\hat{\gamma} \sum_{t=\tau T+1}^T \Delta u_t + 2\hat{\beta} \hat{\gamma} (T - \tau T) \\ T^{-1} \sum_{t=2}^T \Delta \hat{u}_t^2 &\xrightarrow{p} \sigma_\varepsilon^2 [1 + \kappa^2 \{ (1 - \tau_0) + P_2^2 + (1 - \tau) P_3^2 - 2(1 - \tau_0) P_2 - 2(1 - \tau_0 - \mathbb{I}_{\tau_0}^r(\tau - \tau_0)) P_3 \\ &\quad + 2(1 - \tau) P_2 P_3 \}] =: \sigma_\varepsilon^2 \{1 + \kappa^2 L(\tau_0, \tau)\} \end{aligned}$$

and

$$T\hat{\phi} = \frac{T^{-2}\hat{u}_T^2 - T^{-2}\hat{u}_1^2 - T^{-2}\sum_{t=2}^T \Delta\hat{u}_t^2}{2T^{-3}\sum_{t=2}^T \hat{u}_{t-1}^2} \xrightarrow{p} \frac{N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2}{2\int_0^1 N(r, \tau_0, \tau)^2 dr}.$$

Now we can establish the limit of $\hat{\sigma}_\varepsilon^2$

$$\begin{aligned} \hat{\sigma}_\varepsilon^2 &= T^{-1}\sum_{t=2}^T \hat{\varepsilon}_t^2 \\ &= T^{-1}\sum_{t=2}^T \Delta\hat{u}_t^2 + T^2\hat{\phi}^2 T^{-3}\sum_{t=2}^T \hat{u}_{t-1}^2 - 2T\hat{\phi}T^{-2}\sum_{t=2}^T \Delta\hat{u}_t\hat{u}_{t-1} \\ &\xrightarrow{p} \sigma_\varepsilon^2\{1 + \kappa^2 L(\tau_0, \tau)\} + \left[\frac{N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2}{2\int_0^1 N(r, \tau_0, \tau)^2 dr} \right]^2 \kappa^2 \sigma_\varepsilon^2 \int_0^1 N(r, \tau_0, \tau)^2 dr \\ &\quad - 2 \left[\frac{N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2}{2\int_0^1 N(r, \tau_0, \tau)^2 dr} \right] \{\kappa^2 \sigma_\varepsilon^2 N(1, \tau_0, \tau)^2 - \kappa^2 \sigma_\varepsilon^2 N(0, \tau_0, \tau)^2\} \\ &= \sigma_\varepsilon^2 \left(1 + \kappa^2 \left[L(\tau_0, \tau) - \frac{\{N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2\}^2}{2\int_0^1 N(r, \tau_0, \tau)^2 dr} \right] \right) =: \sigma_\varepsilon^2\{1 + \kappa^2 M(\tau_0, \tau)\}. \end{aligned}$$

It then follows using standard theory that

$$T^{-1/2}t_{\hat{\phi}} = \frac{T^{-2}\hat{u}_T^2 - T^{-2}\hat{u}_1^2 - T^{-2}\sum_{t=2}^T \Delta\hat{u}_t^2}{2\sqrt{\hat{\sigma}_\varepsilon^2 T^{-3}\sum_{t=2}^T \hat{u}_{t-1}^2}} \xrightarrow{p} \frac{\kappa^2\{N(1, \tau_0, \tau)^2 - N(0, \tau_0, \tau)^2\}}{2\sqrt{\kappa^2\{1 + \kappa^2 M(\tau_0, \tau)\}\int_0^1 N(r, \tau_0, \tau)^2 dr}}.$$