On Trend Breaks and Initial Condition in Unit Root Testing*

ANTON SKROBOTOV†
The Russian Presidential Academy of National Economy and Public Administration

January 29, 2016

Abstract

Recent approaches in unit root testing have taken into account the influences of initial condition, trend, and breaks in data using pre-testing and union of rejection testing strategies based on obtained information. This paper proposes an extension of the Harvey *et al.* (2012*b*) approach to address the case of uncertainty over the initial condition. It has been shown that this approach has low power under a large initial condition because it includes GLS-based tests. Therefore, the efficiency of some ADF-type unit root tests with breaks under various magnitudes of initial condition will be investigated, and new decision rules will be proposed. Additionally, the modifications of the proposed algorithm, using pre-testing for the trend coefficient and the possible presence of multiple structural trend breaks, are also discussed. The asymptotic behaviors of all tests are analyzed under both a local-to-unity representation of the autoregressive root and a local-to-zero representation of trend and breaks magnitudes. The proposed tests show good asymptotic and finite sample properties under various magnitudes of nuisance parameters.

Key words: unit root test, infimum Dickey-Fuller tests, local trend, local trend break, asymptotic local power, union of rejection, pre-testing, multiple breaks in trend.
JEL: C12, C22.

*Author thanks Robert Taylor for helpful comments.

†Adress correspondence to: Institute of Applied Economic Studies, Russian Presidential Academy of National Economy and Public Administration, office 2103, build 9, 82, Vernadsky pr., 117571, Moscow, Russia.
E-mail: antonskrobotov@gmail.com

1 Introduction

It is well known that in unit root testing it is necessary to take into account the possible presence of structural breaks in the data, and starting with the work of Perron (1989), considerable attention has been paid to the impact of these breaks on unit root testing.

Recent papers, Harvey *et al.* (2012*b*) (hereafter HLT12) and Harvey *et al.* (2013*b*) (hereafter HLT13), address the problem of uncertainty surrounding the presence and dating of structural breaks in the context of unit root testing. An intuitive approach is to use a pre-test to detect the break and then calculate the test statistic with or without this break. However, these methods are only effective in the case of a fixed or zero trend break, in which finite samples produce "valleys" in the power functions of the tests; the power is high for a very small break, but declines rapidly with the increasing magnitude of the break until it increases again. HLT12 proposed two strategies to address this issue. The first strategy recommends always performing tests with the break, but with adaptive critical values. The second approach proposes using the union of rejection of two tests, taking into account scenarios with and without break tests. Additionally, the authors developed a local asymptotic theory for existing and new procedures by using the local-to-zero behavior of the trend break. HLT13 proposed an alternative approach in which the test statistic is computed similarly to Zivot and Andrews (1992) (hereafter ZA), by minimizing the sequence of test statistics for all possible break dates using GLS-detrended data.

However, in the context of unit root testing with an allowance for a break, the issue of the impact of the initial condition is rarely discussed. It can only be found in two studies: Liu and Rodríguez (2006) and Rodrigues (2013). In the former work, the authors developed tests based on GLS-detrending when the initial condition was drawn from the unconditional distribution. In the latter work, the author introduces a test with recursive detrending. The obtained test evidently has lower power than the GLS-based test under zero initial condition, but its power falls much more slowly with an increasing initial condition. This study shows that the OLS-based test (the ZA test) has increasing power with an increasing initial condition. However, the author considers the data generating process with no break. Harvey *et al.* (2013*a*) show that the t-statistic for hypothesis testing for a unit root in an OLS-regression will spuriously reject the null hypothesis with a probability approaching one when the true break fraction is smaller than 2/3 and the break occurs under the null (in contrast to ZA, where no break is present under the null). This demostrates that the ZA test cannot be used for statistical inference. Therefore, in this paper we examine the behavior of the modifications of the ZA proposed by Harvey *et al.* (2013*a*) and Harvey and Leybourne (2012) under various initial conditions and propose algorithms that are robust to initial condition and have high power under small initial condition.

If there are multiple structural breaks, Carrion-i-Silvestre *et al.* (2009) (hereafter CKP) proposed estimating the number of breaks in series in the first step (using a procedure proposed by Kejriwal and Perron (2010)), and then, based on this information, calculating a unit root test statistic, taking into account the number of breaks. This approach, however, as shown in HLT13, results in zero power even with a moderate break magnitude, especially if the break magnitudes have opposite signs. In this paper, we consider possible modifications for solving this problem and discuss their limitations.

The paper is organized as follows: In Section 2 we describe the model with a (local-to-zero) break in trend and seven test statistics considered in the paper. In Section 3 we describe the procedures proposed in HLT12. The investigation of the impact of the initial condition on these tests and the procedures of HLT12 is performed in Section 4. More specifically, we analyze the influence of the initial condition on the so-called robust tests for a trend break in Section 4.1. In Section 4.2, the asymptotic behaviors of unit root tests are compared under various magnitudes of break and initial conditions, and in Section 4.3, we propose the modification of the HLT12 procedures and investigate the asymptotic behaviors of these modifications. Further extensions with the corresponding limitations are considered and discussed in Section 5. Section 6 provides conclusions. The set of Ox programs for calculating all test statistics is available on the author's web page https://sites.google.com/site/antonskrobotov/.

2 The Model

Consider the data generating process (DGP) in the case of a break in trend to be

$$y_t = \mu + \beta t + \gamma_T DT_t(\lambda_0) + u_t,\ t = 1, \dots, T, \tag{1}$$

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

where $DT_t(\lambda_0) = (t - \lfloor \lambda_0 T \rfloor)\mathbb{I}(t > \lfloor \lambda_0 T \rfloor)$, $\mathbb{I}(\cdot)$ is the indicator function, and the trend break occurs at time $\lfloor \lambda_0 T \rfloor$ (λ_0 is the corresponding break fraction) if the break magnitude $\gamma_T \neq 0$. It is assumed that the true break fraction λ_0 is unknown but belongs to the range $\Lambda = [\lambda_L, \lambda_U]$, where $0 < \lambda_L < \lambda_U < 1$, λ_L and λ_U are trimming parameters.[1]

The autoregressive parameter in (2) is taken to be $\rho_T = 1 - c/T$, where $c \geq 0$. Our purpose is testing the null hypothesis of a unit root, $H_0 : \rho_T = 1$ which corresponds to $c = 0$, against the local alternative, $H_1 : \rho_T < 1$ which corresponds to $0 < c < \infty$, without any assumption about whether a break is present in the data or not. We consider the break magnitude as local-to-zero, that is $\gamma_T = \kappa \omega_\varepsilon T^{-1/2}$, as in HLT12 and HLT13, because such a representation provides a better approximation of the finite sample behavior in contrast to the fixed representation $\gamma_T = \gamma$.[2]

A large number of recent papers have investigated the behavior of unit root tests under various initial conditions (see Elliott (1999), Muller and Elliott (2003), Elliott and Muller (2006), Harvey and Leybourne (2005), Harvey and Leybourne (2006) and Harvey *et al.* (2009*b*), *inter alia*). In our paper, in contrast to HLT2012 and HLT2013, we consider asymptotically non-negligible initial conditions according to the following assumption:

Assumption 1 *The initial condition u_1 is defined as $u_1 = \xi = \alpha\sqrt{\omega_\varepsilon^2/(1-\rho_T^2)}$, where $\rho_T = 1 - c/T$, $c > 0$. For $c = 0$, under H_0, the initial condition, without loss of generality, can be set to be zero, $u_1 = 0$, due to the exact similarity of the tests to the initial condition in this case.*

In Assumption 1, α controls the magnitude of the initial condition relative to the innovation long-run variance ω_ε^2. The form given for u_1, allows the initial condition to be either random and of $O_p(T^{1/2})$, or fixed and of $O(T^{1/2})$, depending on whether $\sigma_\alpha^2 > 0$ or $\sigma_\alpha^2 = 0$, respectively.

The linear process ε_t is assumed to satisfy the standard assumptions (see Phillips and Solo (1992)):

[1] It should be noted that in DGP (1)-(2) the trend break is allowed under both the null and alternative hypothesis. Although some tests considered in this study (which are based on the minimization of the test statistics) are constructed under the assumption of the absence of a trend break under the null unit root hypothesis we consider their behavior under non-zero break.

[2] Note that under a fixed magnitude of a break, the results correspond to large values of κ.

Assumption 2 *Let*

$$\varepsilon_t = \gamma(L)e_t = \sum_{i=0}^{\infty} \gamma_i e_{t-i},$$

with $\gamma(z) \neq 0$ *for all* $|z| \leq 1$ *and* $\sum_{i=0}^{\infty} i|\gamma_i| < \infty$*, where* e_t *is the martingale difference sequence with conditional variance* σ_e^2 *and* $\sup_t \mathbb{E}(e_t^4) < \infty$*. The short-run and long-run variances of* ε_t *are defined as* $\sigma_\varepsilon^2 = E(\varepsilon_t^2)$ *and* $\omega_\varepsilon^2 = \lim_{T\to\infty} T^{-1}\mathbb{E}\left(\sum_{t=1}^{T} \varepsilon_t\right)^2 = \sigma_e^2\gamma(1)^2$*, respectively.*

In this paper we analyze the behavior of seven tests. For all tests considered below, the break date is assumed to be unknown.

For the break fraction estimator, we use the hybrid estimator proposed by Harvey and Leybourne (2013):

$$\hat{\lambda}^{D_m} = \arg \min_{\lambda\in\Lambda, \bar{\rho}\in D_m} S(\bar{\rho}, \lambda), \tag{3}$$

where $S(\bar{\rho}, \lambda)$ is the sum of the squared residuals in the regression

$$\boldsymbol{y}^{\bar{\rho}} = \boldsymbol{X}^{\bar{\rho}}(\lambda)\,\beta + \boldsymbol{u}^{\bar{\rho}}, \tag{4}$$

where $\boldsymbol{y}^{\bar{\rho}} = [y_1, (1-\bar{\rho}L)\,y_2, \ldots, (1-\bar{\rho}L)\,y_T]'$, $\boldsymbol{X}^{\bar{\rho}}(\lambda) = [x_1, (1-\bar{\rho}L)\,x_2, \ldots, (1-\bar{\rho}L)\,x_T]'$, $x_t = (1, t, DT_t(\lambda))'$, and $D_m = \{\rho_1', \rho_2', \ldots, \rho_{m-1}', 1\}$ is the m element set, where $|\rho_i'| < 1$ for all i and, without loss of generality, $-1 < \rho_1' < \rho_2' < \cdots < \rho_{m-1}' < 1$. This estimator shows a better properties that the estimator based on the first-differenced regression as in Harris *et al.* (2009) (hereafter HHLT), especially under moderate break magnitude and large initial condition (see Supplementary Appendix, Section 1). Also, the estimator (3) will still be consistent (under a fixed break magnitude), because it use the estimation of the quasi-differencing regressions, see Carrion-i-Silvestre *et al.* (2009). The limiting distribution for $\hat{\lambda}^{D_m}$ follows straightforwardly from HLT12 (Theorem 3 (i, iii)) and continuous mapping theorem (CMT) with GLS-detrending parameters equal to $\bar{c}_\lambda = \{150, 120, 90, 60, 30, 15, 7.5, 3.75, 0\}$ to be consistent with finite sample simulations of Section 4.3, where we use $D_m = \{0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95, 0.975, 1\}$ and $T = 150$.

In the regressions below $\bar{\rho}_T = 1 - \bar{c}/T$ for GLS-based tests without break and $\bar{\rho}_T = 1 - \bar{c}_\lambda/T$ for GLS-based tests with break.[3] In our paper we consider the following tests:

1. The $ADF\text{-}OLS^t$ test is based on the t-statistic for testing $\rho = 1$ in the regression

$$\hat{u}_t = \rho\hat{u}_{t-1} + \sum_{j=1}^{p} \phi_j \Delta\hat{u}_{t-j} + e_t, \quad t = p+2, \ldots, T,$$

   where $\hat{u}_t = y_t - z_t'\hat{\theta}$ are the residuals from the OLS regression of y_t on $z_t = (1, t)'$.

2. The $ADF\text{-}GLS^t$ test is based on the t-statistic for testing $\rho = 1$ in the regression

$$\tilde{u}_t = \rho\tilde{u}_{t-1} + \sum_{j=1}^{p} \phi_j \Delta\tilde{u}_{t-j} + e_t, \quad t = p+2, \ldots, T,$$

   where $\tilde{u}_t$ are the residuals from the OLS regression of $\mathbf{y}_{\bar{c}} = [y_1, y_2 - \bar{\rho}_T y_1, \ldots, y_T - \bar{\rho}_T y_{T-1}]'$ on $\mathbf{Z}_{\bar{c}} = [z_1, z_2 - \bar{\rho}_T z_1, \ldots, z_T - \bar{\rho}_T z_{T-1}]'$, $z_t = (1, t)'$.

[3] We use a single value of $\bar{c}_\lambda$, that does not depend on break fraction estimator for simplicity as in HLT13. See also Section 4.2.

3. The $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ test is based on the t-statistic for testing $\rho = 1$ in the regression

$$\hat{u}_t^{tb} = \rho \hat{u}_{t-1}^{tb} + \sum_{j=1}^{p} \phi_j \Delta \hat{u}_{t-j}^{tb} + e_t, \quad t = p+2, \ldots, T, \tag{5}$$

   where $\hat{u}_t^{tb} = y_t - z_t'\hat{\theta}$ are the residuals from the OLS regression of y_t on $z_t = (1, t, DT_t(\hat{\lambda}^{D_m}))'$.

4. The $ADF\text{-}GLS^{tb}(\hat{\lambda}^{D_m})$ test is based on the t-statistic for testing $\rho = 1$ in the regression

$$\tilde{u}_t^{tb} = \rho \tilde{u}_{t-1}^{tb} + \sum_{j=1}^{p} \phi_j \Delta \tilde{u}_{t-j}^{tb} + e_t, \quad t = p+2, \ldots, T, \tag{6}$$

   where $\tilde{u}_t^{tb}$ are the residuals from the OLS regression $\mathbf{y}_{\bar{c}} = [y_1, y_2 - \bar{\rho}_T y_1, \ldots, y_T - \bar{\rho}_T y_{T-1}]'$ on $\mathbf{Z}_{\bar{c}} = [z_1, z_2 - \bar{\rho}_T z_1, \ldots, z_T - \bar{\rho}_T z_{T-1}]'$, $z_t = (1, t, DT_t(\hat{\lambda}^{D_m}))'$.

5. The $MDF\text{-}OLS_\rho = \inf_{\lambda \in \Lambda} ADF\text{-}OLS_\rho(\lambda)$ test, where $ADF\text{-}OLS_\rho(\lambda)$ is the normalized bias (coefficient) test, $T(\hat{\rho} - 1)/(1 - \sum_{j=1}^{p} \phi_j)$, in the regression (5).

6. The $MDF\text{-}OLS^{\max} = \max(MDF\text{-}OLS, MDF\text{-}OLS')$ test, with

$$MDF\text{-}OLS = \inf_{\lambda \in \Lambda} ADF^{tb}\text{-}OLS(\lambda)$$

   and

$$MDF\text{-}OLS' = \inf_{\lambda \in \Lambda} ADF^{tb}\text{-}OLS(\lambda)',$$

   where $ADF^{tb}\text{-}OLS(\lambda)$ is the t-statistic for testing $\rho = 1$ in regression (5), and $ADF^{tb}\text{-}OLS(\lambda)'$ is also the t-statistic for testing $\rho = 1$ in regression (5), but based on time-reverse data, i.e. the set $\{y_{T-t+1}\}_{t=1}^{T}$ should be used instead of $\{y_t\}_{t=1}^{T}$.

7. The $MDF\text{-}GLS = \inf_{\lambda \in \Lambda} ADF^{tb}\text{-}GLS(\lambda)$ test, where the $ADF^{tb}\text{-}GLS(\lambda)$ is the t-statistic for testing $\rho = 1$ in regression (6).

Lag length is selected by the modified Akaike information criterion (MAIC), proposed by Ng and Perron (2001), with the modification from Perron and Qu (2007).

The first test, $ADF\text{-}OLS^t$, is the Augmented Dickey-Fuller test with trend. The second test, $ADF\text{-}GLS^t$, is the test proposed by Elliott *et al.* (1996), also with trend. In Harvey *et al.* (2009*b*) these two test are compared, and it was determined that the first test is effective under large initial conditions while the second test is effective under small initial conditions, and its power decreases rapidly with an increasing initial condition. The $ADF\text{-}GLS^{tb}(\hat{\lambda}^{D_m})$ test, with a break date estimator based of first differences, was proposed in HHLT, where the authors showed that under fixed break magnitude this test has the same limiting distribution as in the case with a known break date. Similar properties hold for the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ (because of superconsistent break fraction estimator $\hat{\lambda}^{D_m}$, see Kim and Perron (2009)). In Section 4 we compare these tests focusing on the magnitudes of the initial condition and local break in trend. The $MDF\text{-}GLS$ test was proposed by Perron and Rodríguez (2003) and was applied to the case of multiple structural breaks in HLT13. It turned out that the power of this test is more robust to the magnitude of a local break in trend than the other tests considered, even in finite samples.

Lets discuss the OLS-detrended tests, $MDF\text{-}OLS$, $MDF\text{-}OLS_\rho$ and $MDF\text{-}OLS^{\max}$, in greater detail. The first test, the ZA test, would wrongly reject the null hypothesis with a probability approaching unity in the limit when the true break fraction is smaller than 2/3 and the break is present under the null hypothesis (in contrast to ZA, where no break is present under the null hypothesis, see Harvey *et al.* (2013*a*)[4]). For a reasonable sample size and break magnitude, these size distortions occur if the break is in the first half of the sample, and Harvey *et al.* (2013*a*) use this fact. They propose to use the maximum of two statistics, with original and time reversed data, i.e. the $MDF\text{-}OLS^{\max}$ test. The obtained test is robust in the sense that it chooses between $MDF\text{-}OLS$ and $MDF\text{-}OLS'$ on the basis of which statistics are most favorable to the null hypothesis, eliminating the possibility of the size of the test being higher than the nominal level (this, however, holds only in finite samples, not asymptotically). Harvey and Leybourne (2012) considered the asymptotic behavior of the $MDF\text{-}OLS_\rho$ test under a local-to-unit root alternative and also allowing for a local-to-zero break magnitude, $\gamma_T = \kappa\omega_e T^{-1/2}$. The asymptotic size of this test never exceeds the nominal level, except in cases of very small breaks (small κ). In comparison with the asymptotic size of the test based on the conventional t-statistics, the $MDF\text{-}OLS_\rho$ does not have size distortions under a large (fixed) break magnitude, if the true break fraction lies within Λ with $\lambda_L > 0.033$ and standard significance levels are used.

3 Procedures proposed by HLT12

In the HHLT and CKP studies, the following procedures were proposed. To determine whether or not a break is present in the series, the researcher should test its significance with either the Harvey *et al.* (2009*a*) test, t_{HLT}, (or use a modified estimator of the break date according to HHLT) or the Perron and Yabu (2009) test,t_{PY},[5] and then, based on the obtained information about the presence/absence of a break, implement the unit root test with or without a break. When the structural breaks are either large or very small, the power of these procedures is very high. However, considerable power loss is observed in the intermediate range called the power "valley". Power "valleys" occur because, for the range of local break magnitudes, the break is large enough to decrease the power of tests considerably, and, at the same time, is too small to be reliably detected by dating and detection procedures. To the left of this range, the structural breaks are so small that they have no influence on the power of the test without a break. To the right of this range, they are easily detected, so the power increases as the unit root test with a break is performed. Such power "valleys" can be mitigated by always using a unit root test with a break, but its power will be significantly smaller than with the tests from HHLT and CKP in areas with structural breaks that are too small or too large. The authors suggest two alternative procedures that help mitigate the phenomenon of the power "valley".

The first procedure is based on the $ADF\text{-}GLS^{tb}(\tilde{\lambda})$ statistic for unit root testing with GLS-detrended data and an allowance for a break (with a break fraction estimator $\tilde{\lambda}$ as given in (**??**) which minimizes the sum of squared residuals with the first differenced model). Let s_κ denotes a pre-test for a break and cv_κ be the corresponding critical value. Then the adaptive procedure (the

[4] See also Vogelsang and Perron (1998), where the model with innovation breaks is investigated.

[5] We do not provide the exact formulas for the t_{PY} and t_{HLT} tests in order to save space. See a brief description in HLT12, Section 3.

so-called test with adaptive critical values) is

$$A(s_\kappa) = ADF\text{-}GLS^{tb}(\tilde{\lambda}) \text{ with critical value } \begin{cases} cv_{tb}^{consv} & \text{if } s_\kappa < cv_\kappa \\ cv_{tb}^{\tilde{\lambda}} & \text{if } s_\kappa \geq cv_\kappa \end{cases}, \tag{7}$$

where cv_{tb}^{consv} is the conservative critical value, and $cv_{tb}^{\tilde{\lambda}}$ is the critical value for the known break date case obtained in HHLT. In other words, if the structural break is detected, the usual test statistic should be used to test a unit root null hypothesis with a break with the critical value for the known break date, and if the structural break is not detected, the same test statistic is used, but with the conservative critical value cv_{tb}^{consv} (because under zero and small breaks the $ADF\text{-}GLS^{tb}(\tilde{\lambda})$ test will be oversized).

The second procedure, "the adaptive union of rejections of $ADF\text{-}GLS^t$ and $ADF\text{-}GLS^{tb}(\tilde{\lambda})$", attempts to capture some of the power associated with the $ADF\text{-}GLS^t$ when no break is present, while at the same time excluding the possibility of only implementing the $ADF\text{-}GLS^t$ when a break is present. Also, this procedure uses extra power when a break is present in the data (with the help of pre-testing):

$$U(s_\kappa) = \begin{cases} \text{Reject } H_0 \text{ if } \{ADF\text{-}GLS^t < m_\xi cv_t \text{ or} & \\ ADF\text{-}GLS(\tilde{\lambda}) < m_\xi cv_{tb}^{consv}\} & \text{if } s_\kappa < cv_\kappa \\ \text{Reject } H_0 \text{ if } \left\{ADF\text{-}GLS(\tilde{\lambda}) < cv_{tb}^{\tilde{\lambda}}\right\} & \text{if } s_\kappa \geq cv_\kappa \end{cases}. \tag{8}$$

where cv_t is the critical value of the $ADF\text{-}GLS^t$ test and m_ξ is some scaling constant ($m_\xi > 1$) designed to prevent the problem of oversizing and is evaluated under H_0 in cases where there is no break in trend. As the authors' simulations show, the $A(s_\kappa)$ procedure is less sensitive to the magnitude of the break, and the $U(s_\kappa)$ procedure provides much higher power for zero or small breaks due to some power loss for the intermediate magnitudes of a break. Both procedures help to mitigate the effect of power "valleys".

4 The impact of the initial condition

In this section we investigate the asymptotic behavior of all considered tests under various initial conditions. Note that if the initial condition is set as in Assumption 1, then in all limiting distributions under the local alternative the Ornstein-Uhlenbeck process, $W_c(r) = \int_0^r e^{-(r-s)c} dW(s)$ is replaced by

$$K_c(r) = \begin{cases} \alpha(e^{-rc} - 1)(2c)^{-1/2} + W_c(r), & c > 0 \\ W(r), & c = 0 \end{cases}, \tag{9}$$

where $W(r)$ is the standard Wiener process (see Harvey *et al.* (2009*b*)). Results were obtained using simulations of the limiting distributions of test statistics, approximating the Wiener process using $i.i.d.N(0,1)$ random variates and with integrals approximated by normalized sums of 1,000 steps, with 30,000 replications. The limiting distributions for GLS-based tests and also for robust tests can be found in HLT12 and HLT13, under the local behavior of the autoregressive root and the local behavior of the break magnitude, and are not presented here for the sake of brevity. The limiting distribution of the $MDF\text{-}OLS_\rho$ test was obtained from Harvey and Leybourne (2012), the limiting distribution of the $ADF\text{-}OLS^t$ is provided in Lemma 1 of Section 4.2. The limiting distribution of the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ is similar to HLT12 except for the GLS-detrended continuous time residual process should be replaced by corresponding OLS-detrended process (as in Harvey and Leybourne (2012, Theorem 1)). The limiting distribution of $MDF\text{-}OLS^{\max}$ can be obtained in the same manner as in Harvey and Leybourne (2012) under local trend break by applying CMT and arguments provided by Zivot and Andrews (1992). Formal expressions of these results are not presented for the sake of brevity. Also note that we used the trimming parameters λ_L and λ_U, equal to 0.15 and 0.85 respectively. In all simulations, the break fraction is 0.5.

We first investigate the effects of initial conditions on the robust tests from Perron and Yabu (2009) and Harvey *et al.* (2009*a*). We then compare the $MDF\text{-}OLS_\rho$, $MDF\text{-}OLS^{\max}$, $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$, and $ADF\text{-}GLS^{tb}(\hat{\lambda}^{D_m})$, $MDF\text{-}GLS$ tests under various magnitudes of local trend break and initial conditions to determine the effective test in each particular case. After that, we propose modifications to the HLT12 procedures described in Section 3. All necessary critical values and scaling constants are provided in Table 1. The program codes for the simulations and for obtaining critical values are available upon request.

4.1 The influence of the initial condition on the robust tests for trend break

Figures 1-4 show the local power of the t_{HLT} and t_{PY} tests for $c \in \{5, 10, 20, 30\}$ and for $\alpha \in \{-6, -4, -2, -1, 0, 1, 2, 4, 6\}$ with $\kappa \in (0, 10)$. The results with a negative κ are symmetric with respect to α. In general, both tests behave similarly. For a small c, the mean-reversion effect from the initial value is slow enough, so this mean-revering part of the process produces the effect of a breaking trend. This is best seen when $\kappa = 0$ and the test wrongly rejects the hypothesis of no break. Explanation of this phenomenon is similar to Harvey *et al.* (2008). The t_{HLT} and t_{PY} tests use the difference $h \equiv y_T - y_{\lfloor \hat{\lambda} T \rfloor} - (1-\hat{\lambda})(y_T - y_1)$ for estimating the trend break parameter where $\hat{\lambda}$ is the break date estimator equal to argument of supremum of the test statistic under H_0 over all possible break dates (see Harvey *et al.* (2009*a*) for details). Consider the case of $\kappa \geq 0$. If $\kappa = 0$, then the very negative initial condition leads to a large value for $(y_T - y_1)$, which makes h very negative. This leads to a rejection the null hypothesis. If the κ increases, the value of $y_T - y_{\lfloor \hat{\lambda} T \rfloor}$ becomes larger, which make the value of h close to zero, and therefore the test statistic will be insignificant. With a further increase of κ, the value of h increases, which often leads to rejecting the null hypothesis. This result is clearly visible in Figure 1(a). Consider the case of large positive initial condition. Now the value of $(y_T - y_1)$ can be negative, which makes the value of h under $\kappa = 0$ strongly positive. If the κ increases, the rejection rate also increases due to the increases in $y_T - y_{\lfloor \hat{\lambda} T \rfloor}$. Also note that if we compare our results with results of robust tests for trend in Harvey *et al.* (2008), we found that our results are symmetric around α in comparison to Harvey *et al.* (2008), as the value of $(y_T - y_1)$ is included in the numerator of the test statistics with the opposite sign.

For a larger c, the oversizing is less pronounced and often the break is not detected under $\kappa = 0$. This occurs because, for larger c, the larger $|\alpha|$ values would be needed to offset the undersizing in tests when $\alpha = 0$. The negative influence on power in this case is observed only when α and κ have opposite signs. If the signs are the same, then the power is higher than in the case of $\alpha = 0$. Furthermore, for $c = 30$, the tests even have well-controlled sizes (though still conservative).

Thus, although the t_{HLT} and t_{PY} tests cannot be used for a small κ, they can serve as indicators of the large magnitude of a local break even for a large α. In other words, they can be used as pre-tests similar to HLT12.

4.2 The influence of the initial condition on the unit root tests with break

Now consider the behavior of the $MDF\text{-}OLS_\rho$, $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$, $ADF\text{-}GLS^{tb}(\hat{\lambda}^{D_m})$ and $MDF\text{-}GLS$ tests under various κ and α. For illustrating purposes, we also provide the results for the $MDF\text{-}OLS$ test (ZA-test),[6] although this test is invalid under large values of κ.

Figures 5 and 6 show the asymptotic, size-adjusted local power for $c = 20$ and $c = 30$, respectively, and for $\alpha \in \{-6, -4, -2, -1, 0, 1, 2, 4, 6\}$ with $\kappa \in (0, 15)$. More specifically, for the given break fraction λ_0, we calculate the size ($c = 0$) of all tests considered in the previous sections for all $\kappa \in (0, 15)$. Note that the size is not changed for different α, because under H_0 all tests are invariant to initial condition. Let κ^* denote κ, which results in the test's maximum size (the maximum sizes are provided in Table 1). The power curves are obtained by adjusting their size, i.e. by scaling all critical values of the specific test so that its size is equal to 0.05 for $\kappa = \kappa^*$. The same scaling is used for all values of κ (i.e. the size will be always below 0.05 for all $\kappa \neq \kappa^*$).

Also, in contrast to e.g., HHLT, the value of the $\bar{c}_\lambda$ parameter for GLS-based tests is chosen not for each possible location of a break, but as the average of these values (since these values differ very little from each other under the assumption of a local trend break), following HLT13. In other words, we use $\bar{c}_\lambda = 17.6$, proposed in HLT13, for all GLS-based tests.

Based on the figures, it can be seen that for $\alpha = 0$, the $MDF\text{-}GLS$ and $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ are effective. However, for $|\alpha| = 1$, the power curve of the $MDF\text{-}GLS$ is at the same level as the $MDF\text{-}OLS_\rho$, and as the absolute value of α increases, the power of the $MDF\text{-}GLS$ continues to fall. For a large $|\alpha|$, it will be zero, even with a moderate κ (for a κ close to zero, the power of GLS-based tests will be somewhat higher because the mean-reverting effect is interpreted as a break). Similar behavior occurs with the $ADF\text{-}GLS^{tb}(\hat{\lambda}^{D_m})$ test, although the $MDF\text{-}GLS$ is more robust under small and moderate initial conditions.

For moderate and even for large initial conditions, the $ADF\text{-}GLS^{tb}(\hat{\lambda}^{D_m})$ seems to be the somewhat robust. The tests whose power increase by increasing the initial condition are $ADF\text{-}GLS^{tb}(\hat{\lambda}^{D_m})$ and $MDF\text{-}OLS$, but the latter is seriously oversized if the break is occurs in the second part of the sample, and therefore cannot be used. Its modification, $MDF\text{-}OLS^{\max}$, fixes this issue. However, for large $|\alpha|$, the power of the test will be very low because the test includes $MDF\text{-}OLS'$ with time-reversed data, so that the initial value becomes the last value, and the test will rarely reject the null (for a moderate α, this test has reasonably good power, but is strongly dominated by the $MDF\text{-}OLS_\rho$ test. See Supplementary Appendix, Section 1). For the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$, the power decrease are observed for large $|\alpha|$ and small κ, probably due to an incorrect estimation of the break fraction $\hat{\lambda}^{D_m}$ (the mean-reverting effect is estimated as a spurious break). This test would be effective under a large initial condition except for these power decreases for small values of κ. For small breaks, we could use the $MDF\text{-}OLS$ (the ZA test) when the break is small and is not detected by the t_{HLT} and t_{PY} tests. We found that for different break date locations and various types of weak dependence on errors that the $MDF\text{-}OLS$ will be oversized when the break

[6] We omit the results for the $MDF\text{-}OLS^{\max}$ because this test is not effective for any of the the cases examined, and its power falls as the initial condition is increasing.

is detected with a probability of one by the robust tests[7]. In this case the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ could be used because it is correctly sized. If the break is not detected with a nonzero probability, then $MDF\text{-}OLS$ will be correctly sized.

We conclude that, for small initial conditions, the $MDF\text{-}GLS$ should be used, and for large initial conditions, the combination of $MDF\text{-}OLS$ and $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ should be used, dependent on the robust tests for break. Also, for large initial conditions, the test proposed in HLT12 (and also in HLT13) and described in Section 3 has very low power, so its implementation in empirical applications becomes problematic. Furthermore, we propose modifications to these tests with uncertainty over the initial condition.

4.3 The modification of HLT approach

Based on the obtained results, it's obvious that, even for small initial conditions, the power of the HLT12 procedures will fall significantly. Therefore, in this subsection we propose modifications to the $A(s_\kappa)$ and $U(s_\kappa)$ tests that are robust to large initial conditions and at the same time save high power for small initial conditions. In this case, we need a pre-test for the initial condition similar to Harvey *et al.* ($2012a$). However, a pre-test constructed similarly to Harvey *et al.* ($2012a$) as the (weighted) difference between $MDF\text{-}GLS$ and $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ is not appropriate in this context due to power decrease for the latter under small κ. So, we follow the approach of Harvey and Leybourne (2005) and Harvey and Leybourne (2006), and the initial condition estimator, $|\alpha|$, is constructed as

$$|\hat{\alpha}| = \frac{|y_1 - \hat{d}_1|}{\hat{\sigma}}, \tag{10}$$

where $\hat{d}_1$ is the fitted value of the deterministic function in (1) at time $t = 1$, using the break fraction estimator in (3), and $\hat{\sigma}$ is the corresponding standard deviation estimator of the error term u_t. The $|\hat{\alpha}|$ is consistent under a fixed alternative but is not consistent under a local alternative, which we see in the following lemma.

Lemma 1 *Let $\{y_t\}$ be generated as (1) and (2) and Assumptions 1 and 2 are held. Then under $\rho_T = 1 - c/T$, $0 \le c < \infty$*

$$|\hat{\alpha}| \Rightarrow \frac{|K_c^{tb}(0, \lambda_0, \kappa)|}{\sqrt{\int_0^1 K_c^{tb}(r, \lambda_0, \kappa)^2 dr}},$$

where $K_c^{tb}(r, \lambda_0, \kappa)$ is the continuous time residuals process from the regression of $K_c(r) + \kappa(r - \lambda_0)(r > \lambda_0)$ onto the space spanned by $\{1, r, (r - \hat{\lambda}^{D_m})\mathbb{I}(r > \hat{\lambda}^{D_m})\}$, and the limiting distribution for $\hat{\lambda}^{D_m}$ follows straightforwardly from HLT12.

The proof of this lemma comes from Harvey and Leybourne (2005). Note that although the $|\hat{\alpha}|$ is not a consistent estimator for $|\alpha|$ under a local alternative, it can still be useful for obtaining information about a very large $|\alpha|$. In other words, the large value of $|\hat{\alpha}|$ might be associated with large values of $|\alpha|$. We implement the simple heuristic rule that $|\hat{\alpha}| > 1$ indicates a large initial

[7] See Supplementary Appendix, Section 2. In our simulations, for $lambda_0 = 0.2$ the serious size distortions never occur if the break still does not detected, and for $lambda_0 = 0.15$ the size never exceeds 40%. But similar size distortions could occur also for strongly negative MA component in DGP. Also the size distortions decrease if the sample size T increase.

condition. The reason is that for $|\alpha| > 1$, the power of the test that is effective under small initial conditions, $MDF\text{-}GLS$, ceases to be higher than power of other tests.[8] Asymptotic results based on the distribution in Lemma 1 are represented in Figure 7 for various c and κ. It can be seen that this test will be very liberal (and size distortion decreases with an increasing c), but it can still be used under uncertainty over the initial condition for the procedures described below, as a risk-averse strategy. Note that the non-monotonic power for the case of $\kappa = 5$ is a consequence of an incorrect estimation of the break date, as this break is small and the initial condition is large enough.

Now we can consider the modification of the $A(s_\kappa)$ testing strategy, denoted by $A^*(s_\kappa, s_\alpha)$, where, in all algorithms listed below, s_κ and s_α denote tests for break and for initial condition, and cv_κ and cv_α denote corresponding critical values. Critical values and scaling constants are provided in Table 2.

Algorithm 1 *The modified $A^*(s_\kappa, s_\alpha)$ strategy is defined as follows:*

1. *If $s_\kappa \le cv_\kappa$ and $s_\alpha > cv_\alpha$, then use the the $MDF\text{-}OLS$ test with critical values $\delta_\xi \times cv^{MDF\text{-}OLS}$;*

2. *If $s_\kappa \le cv_\kappa$ and $s_\alpha \le cv_\alpha$, then use the strategy*

$$\text{Reject } H_0 \text{ if } \left\{MDF\text{-}OLS < \delta_\xi \times m_\xi^1 \times cv^{MDF\text{-}OLS} \right. \\ \left. \text{or } MDF\text{-}GLS < \delta_\xi \times m_\xi^1 \times cv^{MDF\text{-}GLS}\right\}, \quad (11)$$

   where m_ξ^1 is the scaling constant for the union of the $MDF\text{-}OLS$ and $MDF\text{-}GLS$ tests with conservative critical values.

3. *If $s_\kappa > cv_\kappa$ and $s_\alpha > cv_\alpha$, then use the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ test with the critical values $\delta_\xi \times cv^{ADF\text{-}OLS^{tb}}$, where $cv^{ADF\text{-}OLS^{tb}}$ is the conservative critical value for the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$;*

4. *If $s_\kappa > cv_\kappa$ and $s_\alpha \le cv_\alpha$, then use the strategy*

$$\text{Reject } H_0 \text{ if } \left\{ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m}) < \delta_\xi \times m_\xi^2 \times cv^{ADF\text{-}OLS^{tb}} \right. \\ \left. \text{or } MDF\text{-}GLS < \delta_\xi \times m_\xi^2 \times cv^{MDF\text{-}GLS}\right\}, \quad (12)$$

   where m_ξ^2 is the scaling constant for the union of the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ and $MDF\text{-}GLS$ tests.

In this strategy, δ_ξ is the scaling constant for overall size control.

It should be noted that we use only conservative critical values for all tests calibrated at $\kappa = 0$. The reason is that although the size of tests decreases if κ increases, in finite samples for autocorrelated errors the size is approximately the same for all κ. Thus, the conservative critical values allow better size control.[9]

[8] This decision rule is similar to the rule proposed by Harvey *et al.* (2010) except that Harvey *et al.* (2010) used the trimmed data for estimation of α and the critical value was set to be two. Our approach leads to somewhat better power properties of the initial condition test.

[9] The results for cases with liberal critical values for large κ are available on request.

The basic concepts of this strategy are as follows. Under item 1, there are no reasons to assume that a break is actually present in the data, but there is evidence in favor of a large initial condition. Thus, the $MDF\text{-}OLS$ (ZA) test is effective. Under item 2 there are reasons to assume that both the magnitude of a break and the initial condition are large, so the union of rejection, including the $MDF\text{-}OLS$ and $MDF\text{-}GLS$ tests should be used. Under item 3, there is evidence of both a large magnitude of a break in trend and a large initial condition, so in this case the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ is effective. Under item 4, there is evidence of a large trend break, but there is no reason to believe that the initial condition is large, therefore both the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ and $MDF\text{-}GLS$ should be used.

Now consider the modification of the $U(s_\kappa)$ strategy, denoted by $U^*(s_\kappa, s_\alpha)$. This modification, as well as $U(s_\kappa)$, uses additional tests without breaks to improve the power of the procedure in cases with no breaks. In order to analyze the asymptotic behavior of this strategy, we need, in addition to the known results of HLT12 and HLT13, to obtain the limiting distribution of the $ADF\text{-}OLS^t$ test under a local break in trend, provided in Lemma 2 below.

Lemma 2 *Let* $\{y_t\}$ *be generated as* (1) *and* (2) *and Assumptions 1 and 2 are held. Then under* $\rho_T = 1 - c/T$, $0 \le c < \infty$

$$ADF\text{-}OLS^t \Rightarrow \frac{K_c^t(1, \lambda_0, \kappa)^2 - K_c^t(0, \lambda_0, \kappa)^2 - 1}{2\sqrt{\int_0^1 K_c^t(r, \lambda_0, \kappa)^2\, dr}}, \tag{13}$$

where

$$\begin{aligned} K_c^t(r, \lambda_0, \kappa) = {} & \kappa(r - \lambda_0)\mathbb{I}(r > \lambda_0) - \kappa(1 - \lambda_0)^2/2 \\ & - \kappa(r - 0.5)(1 - 3\lambda_0^2 + 2\lambda_0^3) + K_c^\mu(r) - 12(r - 0.5)\int_0^1 (r - 0.5)K_c(s)ds, \end{aligned} \tag{14}$$

$$K_c^\mu(r) = K_c(r) - \int_0^1 K_c(s)ds,$$

and $K_c(r)$ *is defined in* (9).

The proof of this lemma is similar to the proof of HLT12 for the $ADF\text{-}GLS^t$ test using the results from Harvey *et al.* (2009*b*), and is therefore omitted.

Thus, the $U^*(s_\kappa, s_\alpha)$ strategy is defined in Algorithm 2. Critical values and scaling constants are provided in Table 2.

Algorithm 2 *The modified* $U^*(s_\kappa, s_\alpha)$ *strategy is defined as follows:*

1. *If* $s_\kappa \le cv_\kappa$ *and* $s_\alpha > cv_\alpha$*, then use the strategy*

$$\text{Reject } H_0 \text{ if } \Big\{ ADF\text{-}OLS^t < \tau_\xi \times m_\xi^3 \times cv^{ADF\text{-}OLS^t} \\ \text{or } MDF\text{-}OLS < \tau_\xi \times m_\xi^3 \times cv^{MDF\text{-}OLS} \Big\}, \tag{15}$$

where m_ξ^3 *is the scaling constant for the union of the* $ADF\text{-}OLS^t$ *and* $MDF\text{-}OLS$ *tests with a conservative critical value for the latter.*

2. *If $s_\kappa \leq cv_\kappa$ and $s_\alpha \leq cv_\alpha$, then use the strategy*

$$\text{Reject } H_0 \text{ if } \Big\{ ADF\text{-}OLS^t < \tau_\xi \times m_\xi^4 \times cv^{ADF\text{-}OLS^t} \text{ or } ADF\text{-}GLS^t < \tau_\xi \times m_\xi^4 \times cv^{ADF\text{-}GLS^t} \text{ or } MDF\text{-}OLS < \tau_\xi \times m_\xi^4 \times cv^{MDF\text{-}OLS} \text{ or } MDF\text{-}GLS < \tau_\xi \times m_\xi^4 \times cv^{MDF\text{-}GLS} \Big\}, \quad (16)$$

where m_ξ^4 is the scaling constant for the union of the $ADF\text{-}OLS^t$, $ADF\text{-}GLS^t$, $MDF\text{-}OLS$ and $MDF\text{-}GLS$ tests, with conservative critical values for the latter two.

3. *If $s_\kappa > cv_\kappa$ and $s_\alpha > cv_\alpha$, then use the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ test with the critical values $\tau_\xi \times cv^{ADF\text{-}OLS^{tb}}$;*

4. *If $s_\kappa > cv_\kappa$ and $s_\alpha \leq cv_\alpha$, then use the strategy*

$$\text{Reject } H_0 \text{ if } \Big\{ ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m}) < \tau_\xi \times m_\xi^2 \times cv^{ADF\text{-}OLS^{tb}} \text{ or } MDF\text{-}GLS < \tau_\xi \times m_\xi^2 \times cv^{MDF\text{-}GLS} \Big\}, \quad (17)$$

where m_ξ^2 is the scaling constant for the union of $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ and $MDF\text{-}GLS$.

In this strategy, τ_ξ is the scaling constant for overall size control.

The basic concepts of this strategy are as follows. Under item 1 there are no reasons to assume that a break is actually present in the data, but there is evidence in favor of a large initial condition. Thus, the union of the $MDF\text{-}OLS$ and the $ADF\text{-}OLS^t$ tests should be used. Under item 2, there are no reasons to assume that both the magnitude of the break and the initial condition are large, so the union of rejection, including four tests, the $ADF\text{-}OLS^t$, $ADF\text{-}GLS^t$, $MDF\text{-}OLS$ and $MDF\text{-}GLS$ should be used. The 3rd and 4th items are the same as the 3rd and 4th items from the $A^*(s_\kappa, s_\alpha)$ strategy except for the other overall scaling constant.

Figures 8 and 9 show the asymptotic size-adjusted power (the correction is performed similarly to the way it was performed in the previous subsection, the maximum sizes across κ are provided in Table 1) for $c = 20$ and $c = 30$, respectively.[10] These results correspond to the behavior of the unit root tests and pre-tests and show the robustness of the proposed strategies for various initial conditions. The $A^*(t_{HLT}, s_\alpha)$ strategy seems to be the most robust because under small initial condition the power curve lies between the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ and the effective $MDF\text{-}GLS$, under large initial condition and small break this strategy considerably improve the power "valley" effect of the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ test, and under large initial condition and large break the power close to the effective $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$. The $U^*(t_{HLT}, s_\alpha)$ strategy still shows small power "valleys" for small α, although somewhat outperform all tests for small κ.

We also consider the behavior of the tests for DGP (1)-(2) in finite samples when the error terms are generated according to either an $i.i.d$ sequence or the AR(1) and MA(1) processes with the sample size $T = 150$. This corresponds to $\varepsilon_t \sim i.i.d.N(0,1)$ for the case of $i.i.d.$, $\varepsilon_t = 0.5\varepsilon_{t-1} + e_t$ for the case of AR(1), and $\varepsilon_t = e_t - 0.5e_{t-1}$ for the case of MA(1), where $e_t \sim i.i.d.N(0,1)$. For break date estimator (3) we use $D_m = \{0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95, 0.975, 1\}$. The maximum sizes across κ are provided in Table 1. For $i.i.d$ and AR(1) cases the size is close to the nominal

[10] All results regarding the power of the tests are provided only for the t_{HLT} for brevity.

one, for MA(1) case the liberal size distortion is sufficiently large but this is standard results (see the corresponding tables in HLT12 and HLT13). The results are presented in Figures 8 and 9. Qualitatively, the results are consistent with the asymptotic, especially for $i.i.d.$ case. However, the $A^*(t_{HLT}, s_\alpha)$ still remains the most robust strategy that we recommend to use in practice.

5 Further extensions

Based on the properties of our proposed modifications described in the previous section, these procedures can be expanded in several directions. Furthermore, we briefly consider the following generalizations indicating the associated limitations: using additional information from the test for trend, the presence of more than one structural break, and the presence of *a priori* information about the location of the break.

5.1 Uncertainty over the trend

In some situations an additional uncertainty can exist. We restrict our model to the condition that if there is a break in trend, then the coefficient of the trend is exactly non-zero (in fact, if the break is present, the test statistics are independent of the trend magnitude). However, if there is no trend break, it is possible that there is no trend in the data, and in this case, only the tests that take the constant in their construction into account are effective. In other words, the $ADF\text{-}OLS^\mu$ and $ADF\text{-}GLS^\mu$ tests (which are constructed in a similar way to $ADF\text{-}OLS^t$ and $ADF\text{-}GLS^t$, but without trend) should be included in our $U^*(s_\kappa, s_\alpha)$ strategy. It would be possible to generalize Algorithm 2 by additionally including the $ADF\text{-}OLS^\mu$ and $ADF\text{-}GLS^\mu$ tests if (sequentially) first the break is not detected, and then the trend is not detected.

However, when using such a procedure, one may encounter two problems. The first is that the power of tests for trend will be very low when there is a small undetectable break with the opposite sign. The second problem is that when both a break and trend are undetected, the strategy should include six tests, so the critical values should be multiplied by a scaling constant of 1.3 (at the 5 percent significance level) to control the size in the absence of a break and trend. This will have a negative effect on the power. More detailed results from the simulations are available on request.

5.2 Multiple structural breaks

The test procedures discussed above can be used in a case of multiple structural breaks, because if the number of breaks is more than taken into account when constructing the test, the power will drop to zero. For the unit root tests we can consider two types of tests based on OLS- or GLS-detrending, respectively, similar to the previous section. The infimum test based on GLS-detrending was proposed by HLT2013 in the context of a case of multiple breaks, and this test was effective under small initial conditions. The extensions of the OLS-based tests are constructed in a similar way.

However, the problem of the construction of the effective combination of OLS-based tests under large initial conditions has arisen. The behavior of the $MDF\text{-}OLS$ with multiple breaks is not investigated under fixed break magnitudes and unclear how large the magnitudes of breaks should be in order experience serious size distortions with $MDF\text{-}OLS$. If we find an effective combination of OLS-based tests, it could be used in similar strategies to those described in Section 4, with the Kejriwal and Perron (2010) and Sobreira and Nunes (2012) procedures used as pre-tests to determine the number of breaks (see also Sobreira *et al.* (2014)). However, the power "valleys" for these strategies can be much more severe and can increase with the number of tests. One might ask whether it makes sense in the case of, for example, three breaks, to use a fairly complex combination of tests, and the best solution is to use only one test (with the maximum number of breaks). We have left all these questions open for future research. Currently, the test based on (augmented) Dickey-Fuller regression with with break dates estimators according Harvey and Leybourne (2013) (the $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$ in case of multiple breaks) seems to be the most robust.

5.3 *A priori* information on the break date

One modification when testing with the allowance for one structural break was proposed in Harvey *et al.* (2014). Many unit root tests with a break use the fact that the break can occur in any location, except for some values at the beginning and end of the series. However, the researcher may know some *a priori* information about the location of the break without knowing its exact location. This approach was first considered by Andrews (1993) when testing for general structural instability by motivating it using two examples: a significant political event (economic reform, war, etc.) can occur in a certain period of time, but it is unknown exactly when the effects begin; the event may occur on a certain date, but the effect occurs with some delay.

A priori information about the date of the break implies that the true break fraction is $\lambda_0 \in \Lambda(\tau_{mid}, \delta)$, where $\Lambda(\tau_{mid}, \delta)$ is a window in which the break occurs. This window is defined as $\Lambda(\tau_{mid}, \delta) = [\tau_{mid} - \delta/2, \tau_{mid} + \delta/2]$, where $\delta > 0$ denotes the width of the window containing all permissible break fractions, and τ_{mid} denotes the mid-point of the window. $\tau_{mid} - \delta/2 > 0$ and $\tau_{mid} + \delta/2 < 1$ is a requirement.

Therefore, all minimizations considered in this paper are performed not on the whole set $\Lambda = [\lambda_L, \lambda_U]$, but on the set $\Lambda(\tau_{mid}, \delta)$. It is clear that the power of the test increases when there is more known *a priori* information about the location of the true break date, i.e. with a decreasing value of δ. The location of the break within the selected window has a small effect on the asymptotic size and power. However, the size and power will be seriously downward biased if the information about the true break date location is wrong, and the distortion increases as the error of this information increases. Thus, all procedures considered in this paper can be easily expanded to cases of partial information about the break date location.[11] The ox-code for calculating the critical values and scaling constants for different δ and τ_{mid} is available on the author's web-page.

6 Conclusion

In this paper different types of uncertainty concerning trend, break, and initial conditions were analyzed. Under various magnitudes of the trend/break/initial condition, which are usually unknown *a priori*, different tests are effective. It was shown that GLS-based unit root tests with a break had low power under large initial conditions, much like the conventional GLS-based test without a break. However, they have maximum power across all considered tests under zero initial

[11] Note, that the $MDF\text{-}OLS$ test can be used for all κ if the break is occurred in the second half of sample.

condition. The behavior of different tests with a break based on OLS-detrending was analyzed, and algorithms allowing for uncertainty concerning both the trend break and the initial condition were proposed. Thus, the proposed algorithms are useful in empirical applications, because, in contrast to existing approaches, we examined the situation of simultaneous uncertainty about trend breaks and initial value.

Also we discussed possible modifications of the proposed strategies: allowing additional uncertainty over the trend and using the test with the absence of break, the possible presence of multiple breaks, and using partial information about break date location.

References

Andrews, D.W.K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica*, **61**, 821–856.

Carrion-i-Silvestre, J.L., Kim, D., and Perron, P. (2009). GLS-Based Unit Root Tests with Multiple Structural Breaks Both under the Null and the Alternative Hypotheses. *Econometric Theory*, **25**, 1754–1792.

Elliott, G. (1999). Efficient Tests for a Unit Root when the Initial Observation is Drawn from its Unconditional Distribution. *International Economic Review*, **40**, 767–783.

Elliott, G. and Muller, U. (2006). Minimizing the Impact of the Initial Condition on Testing for Unit Roots. *Journal of Econometrics*, **135**, 285–310.

Elliott, G., Rothenberg, T.J., and Stock, J.H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, **64**, 813–836.

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

Harvey, D.I. and Leybourne, S.J. (2005). On Testing for Unit Roots and the Initial Observation. *The Econometrics Journal*, **8**, 97–111.

Harvey, D.I. and Leybourne, S.J. (2006). Power of a Unit-Root Test and the Initial Condition. *Journal of Time Series Analysis*, **27**, 739–752.

Harvey, D.I. and Leybourne, S.J. (2012). An Infimum Normalized Bias Unit Root Test Allowing for an Unknown Break in Trend. *Economics Letters*, **117**, 298–302.

Harvey, D.I. and Leybourne, S.J. (2013). Break Date Estimation for Models with Deterministic Structural Change. *Oxford Bulletin of Economics and Statistics*, **76**, 623–642.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2008). Testing for Unit Roots in the Presence of Uncertainty over Both the Trend and Initial Condition. Granger Centre Discussion Paper No. 08/03, University of Nottingham.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2009*a*). Simple, Robust and Powerful Tests of the Breaking Trend Hypothesis. *Econometric Theory*, **25**, 995–1029.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2009*b*). Unit Root Testing in Practice: Dealing with Uncertainty over the Trend and Initial Condition (with Commentaries and Rejoinder). *Econometric Theory*, **25**, 587–667.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2010). The Impact of the Initial Condition on Robust Tests for a Linear Trend. *Journal of Time Series Analysis*, **31**, 292–302.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2012*a*). Testing for Unit Roots in the Presence of Uncertainty over Both the Trend and Initial Condition. *Journal of Econometrics*, **169**, 188–195.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2012*b*). Unit Root Testing under a Local Break in Trend. *Journal of Econometrics*, **167**, 140–167.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2013*a*). On Infimum Dickey-Fuller Unit Root Tests Allowing for a Trend Break under the Null. *Computational Statistics and Data Analysis*, **78**, 235–242.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2013*b*). Testing for Unit Roots in the Possible Presence of Multiple Trend Breaks using Minimum Dickey-Fuller Statistics. *Journal of Econometrics*, **177**, 265–284.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2014). Unit Root Testing under a Local Break in Trend using Partial Information on the Break Date. *Oxford Bulletin of Economics and Statistics*, **76**, 93–111.

Kejriwal, M. and Perron, P. (2010). A Sequential Procedure to Determine the Number of Breaks in Trend with an Integrated or Stationary Noise Component. *Journal of Time Series Analysis*, **31**, 305–328.

Kim, D. 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.

Liu, H. and Rodríguez, G.H. (2006). Unit Root Tests and Structural Change when the Initial Observation is Drawn from its Unconditional Distribution. *Econometrics Journal*, **9**, 225–251.

Muller, U. and Elliott, G. (2003). Tests for Unit Roots and the Initial Condition. *Econometrica*, **71**, 1269–1286.

Ng, S. and Perron, P. (2001). Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica*, **69**, 1519–1554.

Perron, P. (1989). The Great Crash, the Oil Price Shock and the Unit Root Hypothesis. *Econometrica*, **57**, 1361–1401.

Perron, P. and Qu, Z. (2007). A Simple Modification to Improve the Finite Sample Properties of Ng and Perron's Unit Root Tests. *Economics Letters*, **94**, 12–19.

Perron, P. and Rodríguez, G.H. (2003). GLS Detrending, Efficient Unit Root Tests and Structural Change. *Journal of Econometrics*, **115**, 1–27.

Perron, P. and Yabu, T. (2009). Testing for Shifts in Trend with an Integrated or Stationary Noise Component. *Journal of Business and Economic Statistics*, **27**, 369–396.

Phillips, P.C.B. and Solo, V. (1992). Asymptotics for Linear Processes. *Annals of Statistics*, **20**, 971–1001.

Rodrigues, P.M.M. (2013). Recursive Adjustment, Unit Root Tests and Structural Breaks. *Journal of Time Series Analysis*, **34**, 62–82.

Sobreira, N. and Nunes, L.C. (2012). Tests for Multiple Breaks in the Trend with Stationary or Integrated Shocks. Unpublished Manuscript.

Sobreira, N., Nunes, L.C., and Rodrigues, P.M.M. (2014). Characterizing economic growth paths based on new structural change tests. *Economic Inquiry*, **52**, 845–861.

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.

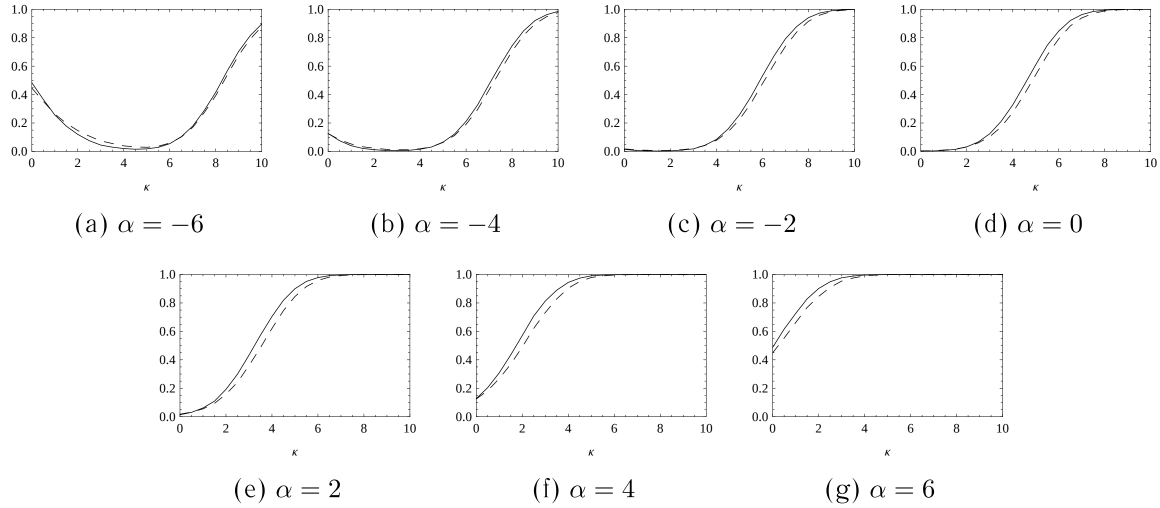
Zivot, E. and Andrews, D.W.K. (1992). Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis. *Journal of Business and Economic Statistics*, **10**, 251–270.

Table 1. Maximum finite sample sizes of nominal 0.05-level tests across $\kappa \in (0, 16)$

	$T = \infty$	$i.i.d.$ errors	$AR(1)$ errors	$MA(1)$ errors
$MDF\text{-}GLS$	0.051	0.051	0.057	0.111
$ADF\text{-}OLS(\hat{\lambda}^{D_m})$	0.053	0.063	0.042	0.087
$MDF\text{-}OLS$	0.052	0.048	0.050	0.099
$MDF\text{-}OLS_\rho$	0.053	0.032	0.077	0.254
$U^*(t_{HLT}, s_\alpha)$	0.053	0.063	0.047	0.102
$U^*(t_{PY}, s_\alpha)$	0.053	0.056	0.048	0.101
$A^*(t_{HLT}, s_\alpha)$	0.053	0.067	0.059	0.115
$A^*(t_{PY}, s_\alpha)$	0.053	0.065	0.060	0.115

Table 2. Asymptotic critical values and scaling constants at ξ-level

ξ	0.01	0.05	0.10
$cv^{ADF\text{-}GLS^t}$	-3.41	-2.84	-2.54
$cv^{ADF\text{-}OLS^t}$	-3.95	-3.40	-3.11
$cv^{MDF\text{-}GLS}$	-4.37	-3.85	-3.56
$cv^{MDF\text{-}OLS}$	-4.79	-4.25	-3.99
$cv^{ADF\text{-}OLS^{tb}}$	-4.60	-4.08	-3.80
m_ξ^1	1.03	1.04	1.04
m_ξ^2	1.03	1.03	1.03
m_ξ^3	1.04	1.05	1.06
m_ξ^4	1.07	1.10	1.12
$\tau_\xi(t_{HLT})$	1.02	1.02	1.02
$\delta_\xi(t_{HLT})$	1.00	1.00	1.00
$\tau_\xi(t_{PY})$	1.02	1.02	1.02
$\delta_\xi(t_{PY})$	1.00	1.00	1.00

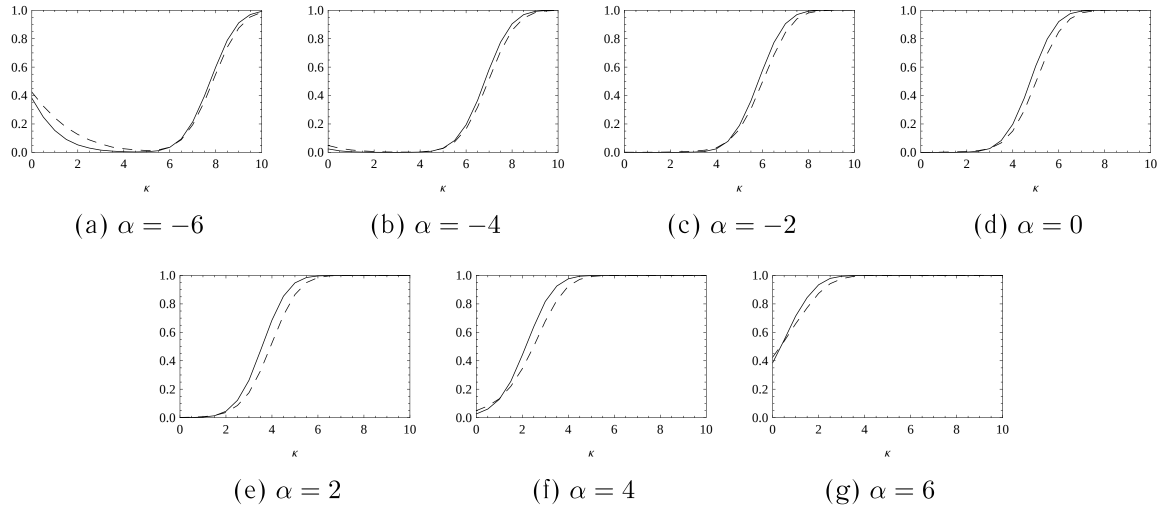


(a) $\alpha = -6$ (b) $\alpha = -4$ (c) $\alpha = -2$ (d) $\alpha = 0$

(e) $\alpha = 2$ (f) $\alpha = 4$ (g) $\alpha = 6$

Figure 1. Asymptotic local power of s_κ, $c = 5$

t_{HLT} : —— , t_{PY} : — —

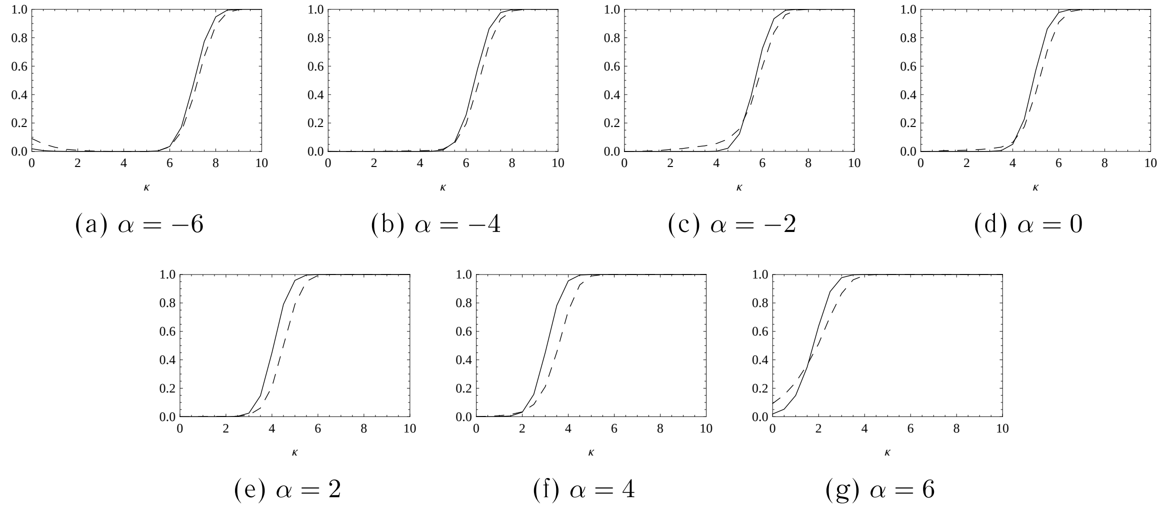


(a) $\alpha = -6$ (b) $\alpha = -4$ (c) $\alpha = -2$ (d) $\alpha = 0$

(e) $\alpha = 2$ (f) $\alpha = 4$ (g) $\alpha = 6$

Figure 2. Asymptotic local power of s_κ, $c = 10$

t_{HLT} : —— , t_{PY} : — —

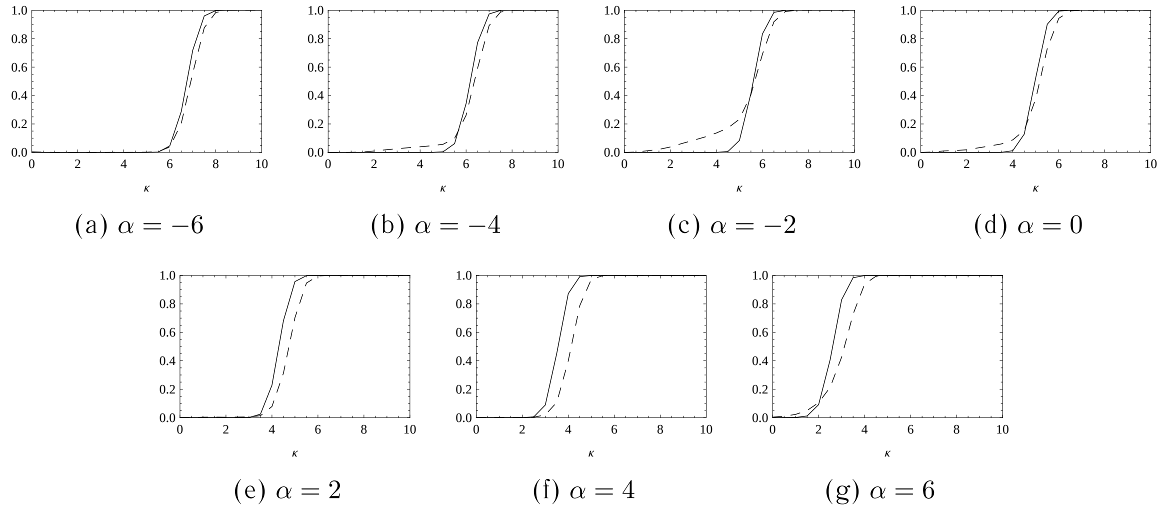


(a) $\alpha = -6$ (b) $\alpha = -4$ (c) $\alpha = -2$ (d) $\alpha = 0$

(e) $\alpha = 2$ (f) $\alpha = 4$ (g) $\alpha = 6$

Figure 3. Asymptotic local power of s_κ, $c = 20$

t_{HLT} : ——— , t_{PY} : — —

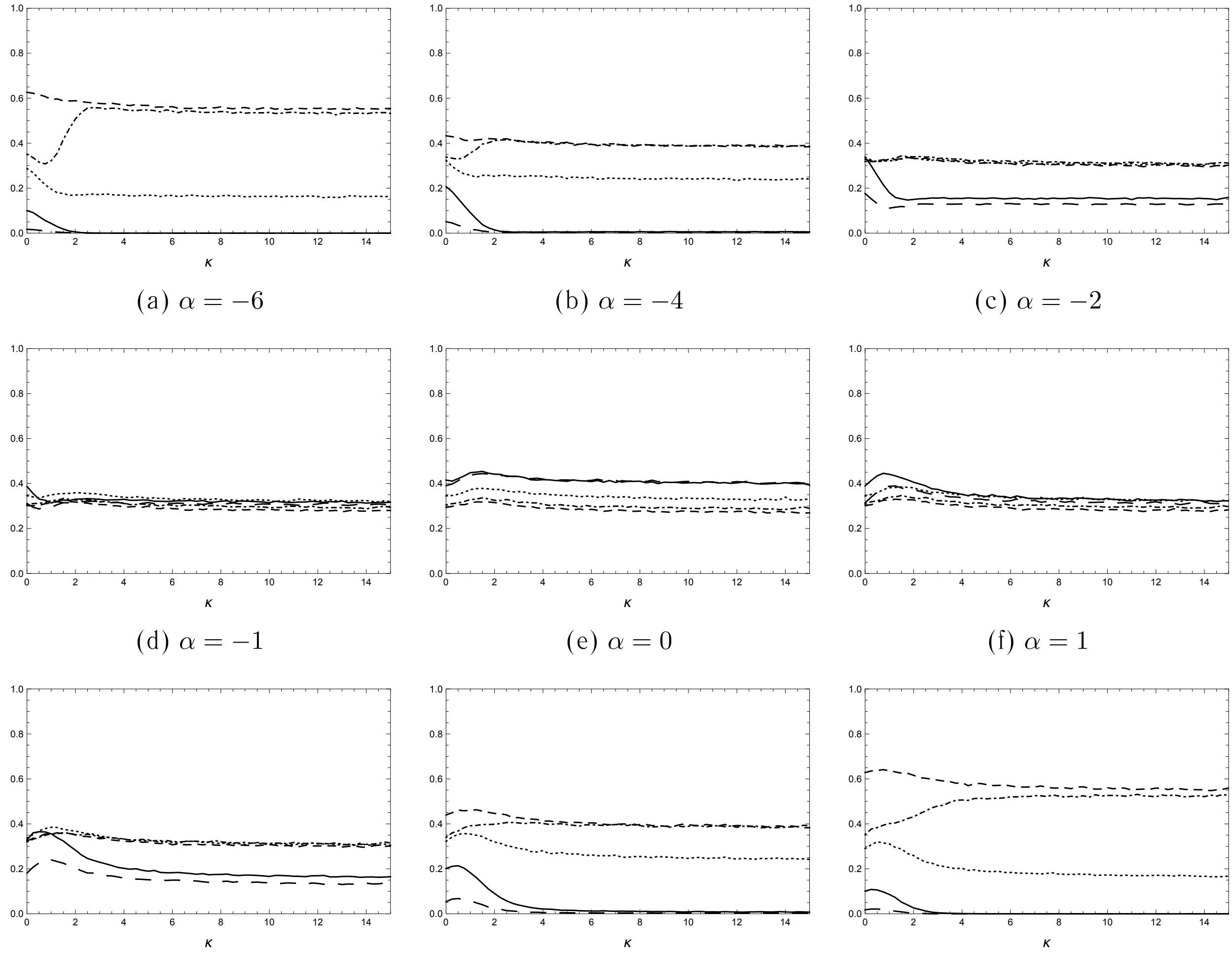


(a) $\alpha = -6$ (b) $\alpha = -4$ (c) $\alpha = -2$ (d) $\alpha = 0$

(e) $\alpha = 2$ (f) $\alpha = 4$ (g) $\alpha = 6$

Figure 4. Asymptotic local power of s_κ, $c = 30$

t_{HLT} : ——— , t_{PY} : — —



(a) $\alpha = -6$ (b) $\alpha = -4$ (c) $\alpha = -2$

(d) $\alpha = -1$ (e) $\alpha = 0$ (f) $\alpha = 1$

(g) $\alpha = 2$ (h) $\alpha = 4$ (i) $\alpha = 6$

Figure 5. Asymptotic local power, $c = 20$

$MDF\text{-}GLS$: ——— , $ADF\text{-}GLS(\hat{\lambda}^{D_m})$: — — , $MDF\text{-}OLS$: - - , $MDF\text{-}OLS_\rho$: · · · ,
$ADF\text{-}OLS(\hat{\lambda}^{D_m})$: **— · —**

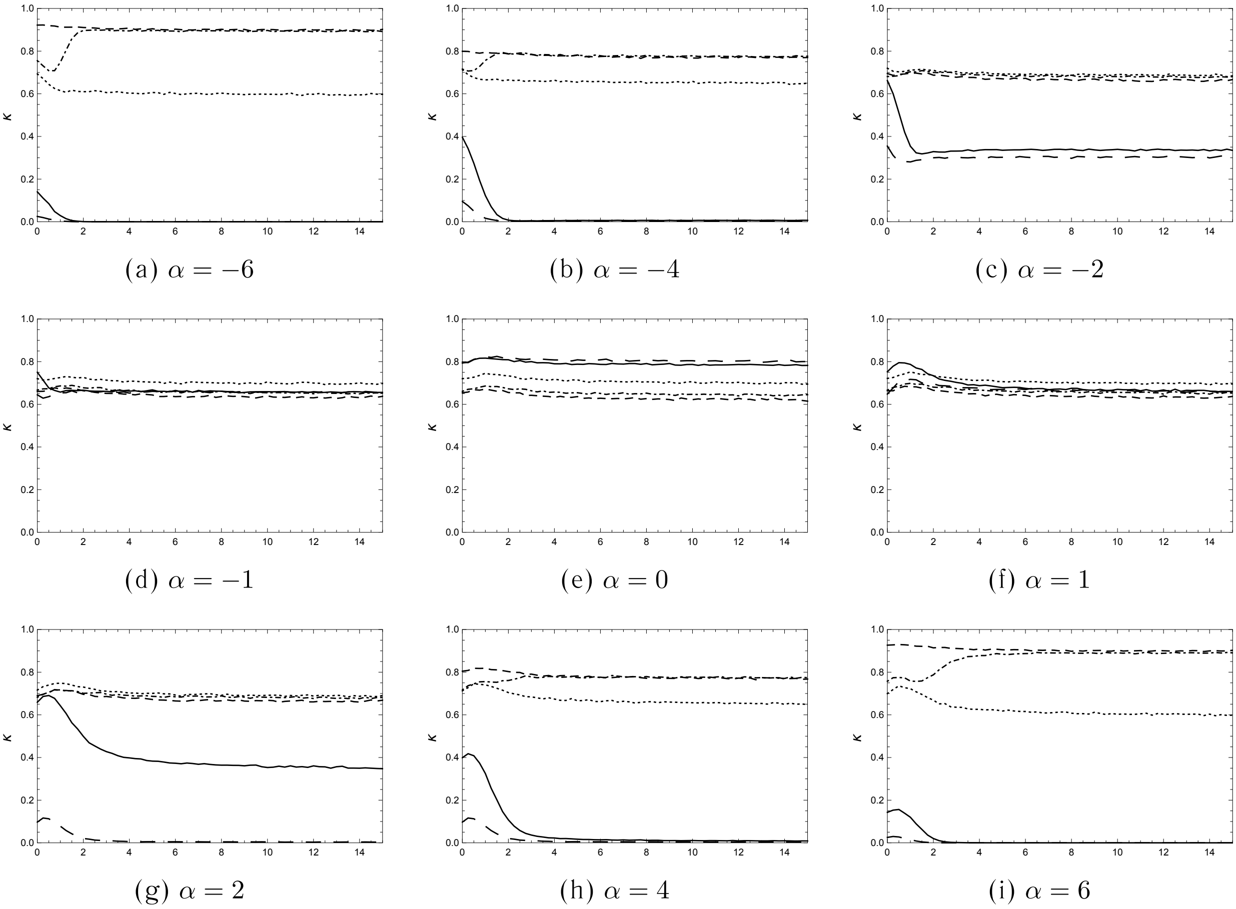
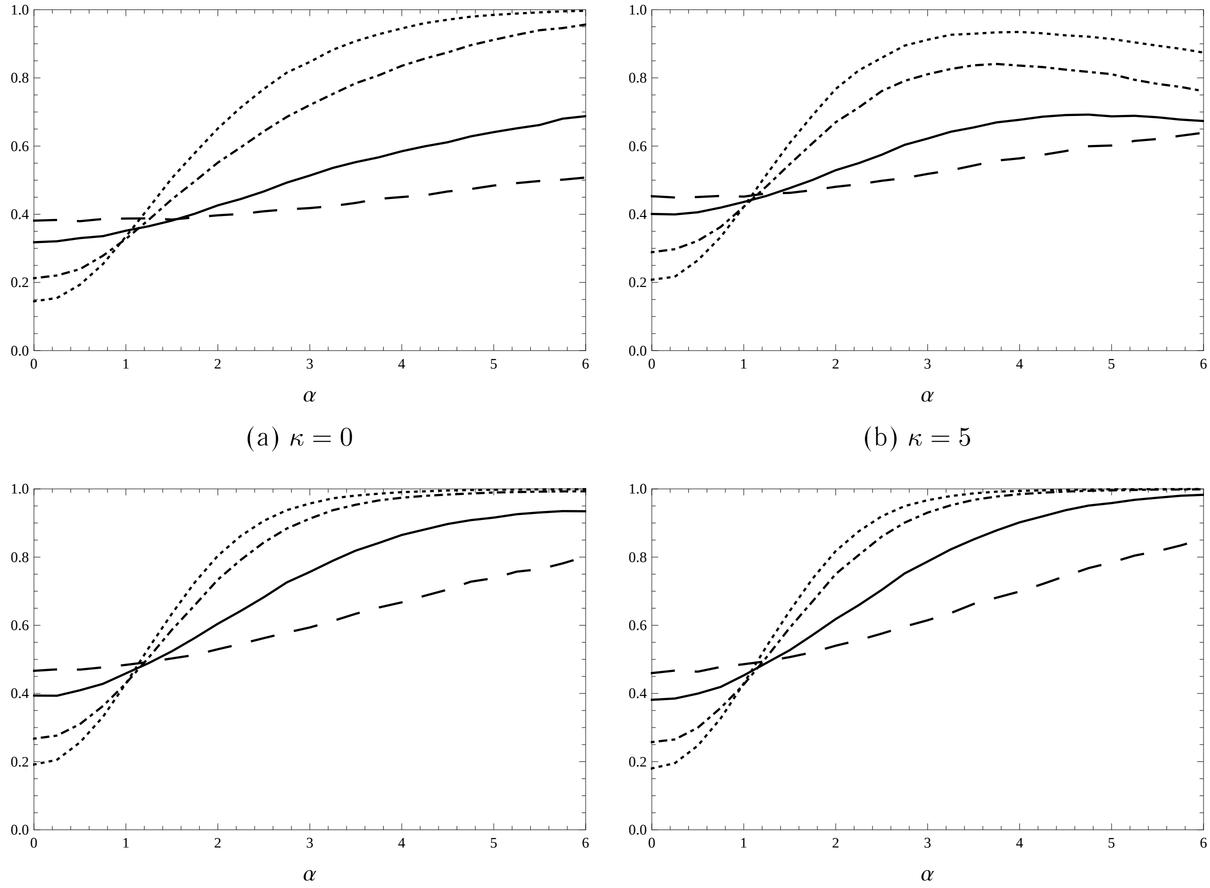


Figure 6. Asymptotic local power, $c = 30$

$MDF\text{-}GLS$: ——, $ADF\text{-}GLS(\hat{\lambda}^{D_m})$: — —, $MDF\text{-}OLS$: - -, $MDF\text{-}OLS_\rho$: · · · , $ADF\text{-}OLS(\hat{\lambda}^{D_m})$: —·—

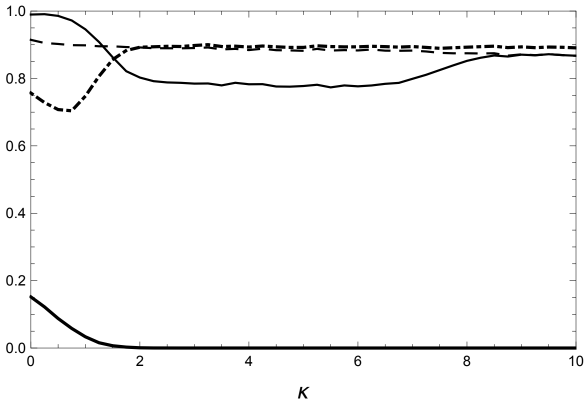


(a) $\kappa = 0$ (b) $\kappa = 5$

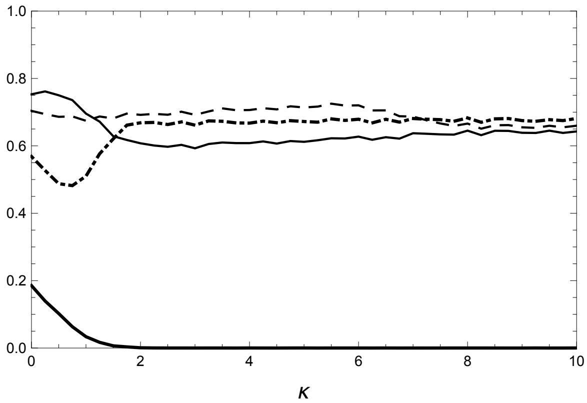
(c) $\kappa = 10$ (d) $\kappa = 15$

Figure 7. Asymptotic local power, s_α

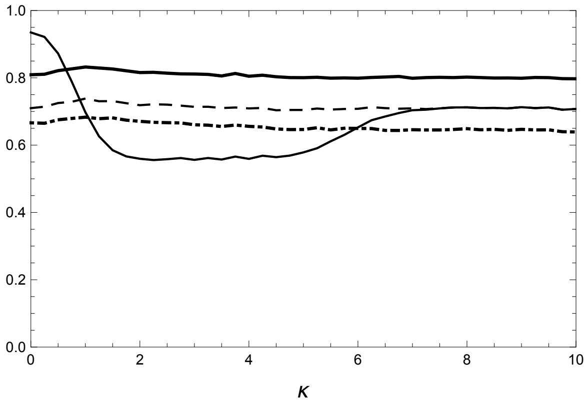
$c = 5$: — —, $c = 10$: ——, $c = 20$: — · — $c = 30$: · · ·



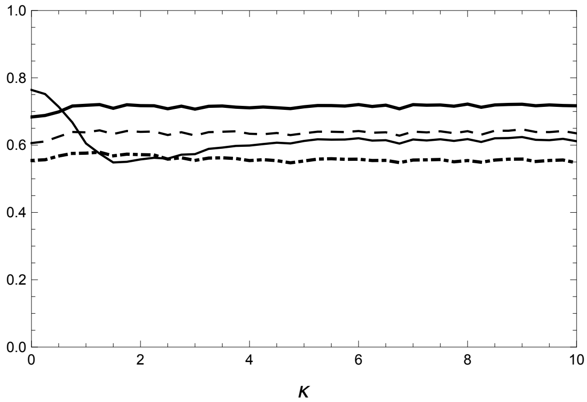
(a) $\alpha = -6$



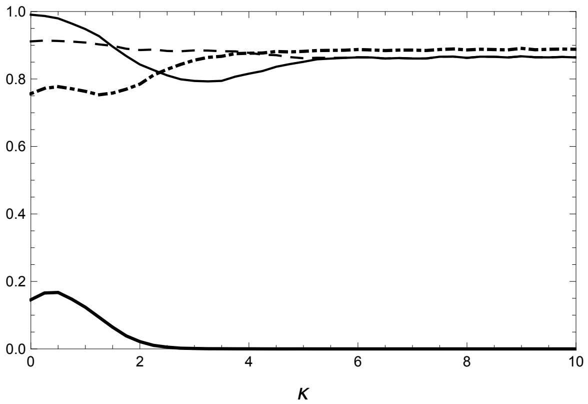
(b) $\alpha = -6$



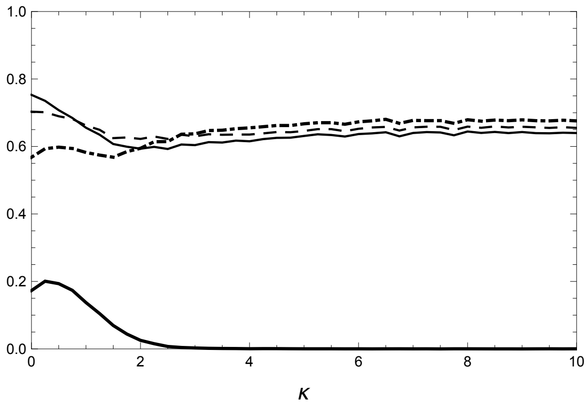
(c) $\alpha = 0$



(d) $\alpha = 0$

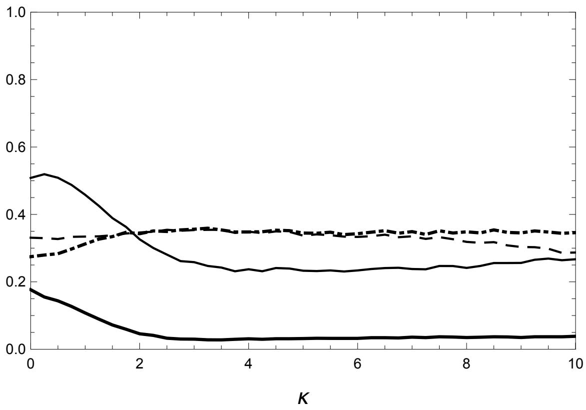


(e) $\alpha = 6$

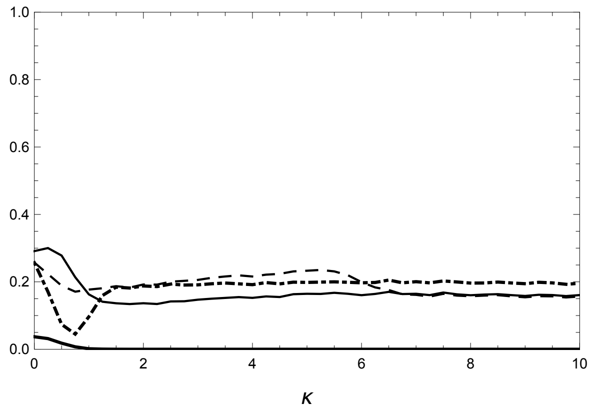


(f) $\alpha = 6$

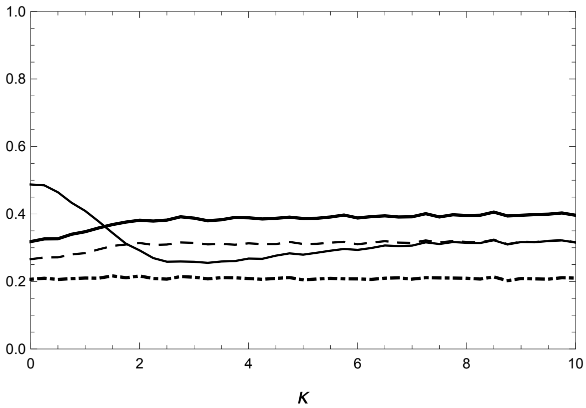
Figure 8. Asymptotic local power (left) and finite sample power with $i.i.d.$ errors (right), $c = 30$
$MDF\text{-}GLS$: ▬▬ , $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$: ▬ · ▬ $UR^*(t_{HLT}, s_\alpha)$: —— , $A^*(t_{HLT}, s_\alpha)$: – – ,



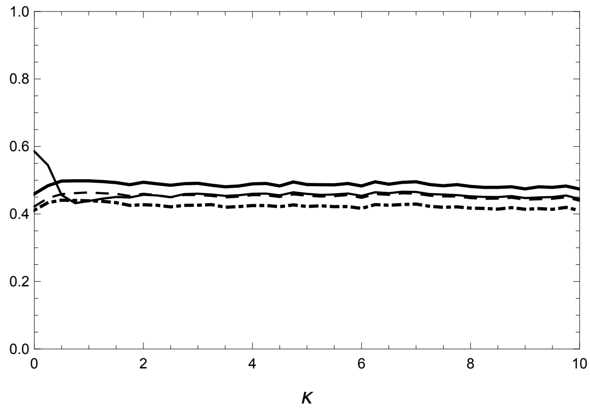
(a) $\alpha = -6$



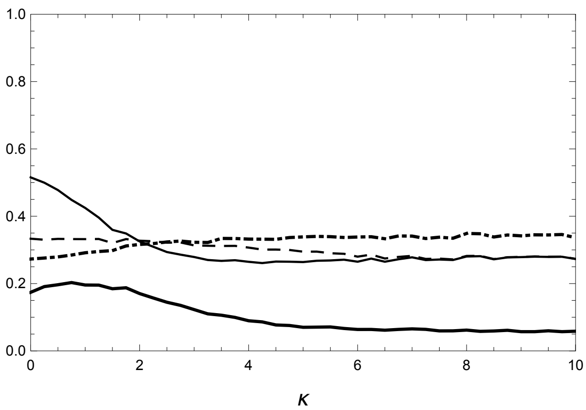
(b) $\alpha = -6$



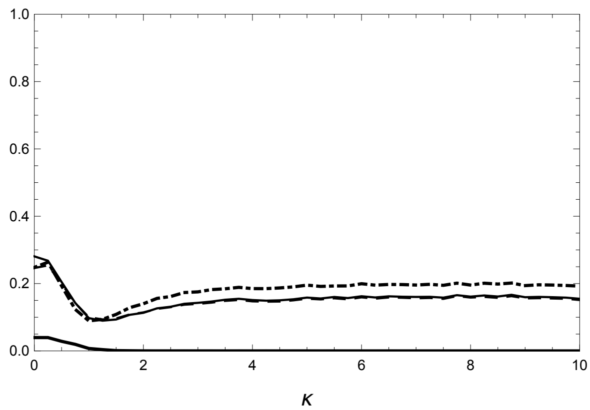
(c) $\alpha = 0$



(d) $\alpha = 0$



(e) $\alpha = 6$



(f) $\alpha = 6$

Figure 9. Finite sample power with $AR(1)$ errors (left) and finite sample power with $MA(1)$ errors (right), $c = 30$

$MDF\text{-}GLS$: ▬▬ , $ADF\text{-}OLS^{tb}(\hat{\lambda}^{D_m})$: ▬ · ▬ $UR^*(t_{HLT}, s_\alpha)$: —— , $A^*(t_{HLT}, s_\alpha)$: - - ,