Trend and initial condition in stationarity tests: the asymptotic analysis*

Anton Skrobotov†
Gaidar Institute for Economic Policy,
The Russian Presidential Academy
of National Economy and Public Administration

October 1, 2013

Abstract

In this paper we investigate the behavior of stationarity tests proposed by Müller (2005) and Harris *et al.* (2007) with uncertainty over the trend and/or initial condition. As different tests are efficient for different magnitudes of local trend and initial condition, following Harvey *et al.* (2012) we propose decision rule based on the rejection of null hypothesis for multiple tests. Additionally, we propose a modification of this decision rule, relying on additional information about the magnitudes of the local trend and/or the initial condition that is obtained through pre-testing. The resulting modification has satisfactory size properties under both uncertainty types.
Key words: Stationarity test, KPSS test, uncertainty over the trend, uncertainty over the initial condition, size distortion, intersection of rejection decision rule.
JEL: C12, C22

1 Introduction

The influence of linear trend and/or the initial condition can be very important in unit root testing. In recent works, Harvey *et al.* (2009) and Harvey *et al.* (2012) (hereafter HLT, see also Harvey *et al.* (2008)) considered the issue of deterministic trend inclusion in the unit root test and investigated the behavior of tests with different initial conditions. Harvey *et al.* (2009) showed that under uncertainty over the linear trend, the optimal unit root test is a simple union of rejections of the two tests (i.e., the null hypothesis of the unit root is rejected, if it is rejected by at least one of the tests): the first with inclusion of the linear trend, the second only with the constant. Both of these tests need to be effective for their respective types of deterministic components in the absence of large initial condition. Simultaneously, knowing the type of deterministic components (i.e. whether the trend is present in the data or not) and uncertainty over the initial condition, the union of rejection

*We thank two anonymous referees, Vladimir Nosko and Marina Turuntseva for helpful comments and recommendations for improvement earlier version of this paper.
†E-mail: antonskrobotov@gmail.com

testing strategy of two tests (one effective at a small initial condition, and the second effective at a large initial condition for a given type of deterministic component) will be the best. HLT extended the procedure by assuming uncertainty over both the trend and initial condition and suggested a union of rejection testing strategy for all four tests. Additionally, HLT also suggested a modification of this test with pre-testing of the linear trend coefficient and the initial condition. Then, if there was evidence of a large local trend and/or evidence of a large initial condition, it was more likely that a trend and/or a large initial condition were actually present in the data. Thus, this information can be used to construct the union of rejection testing strategy.

There is a need of a similar procedure for stationarity tests, as hypothesis testing opposed to unit root is important for confirmatory analysis (see, e.g., Maddala and Kim (1998, Ch. 4.6)). Harris *et al.* (2007) (hereafter HLM) proposed a modification of the standard Kwiatkowski *et al.* (1992) test (hereafter KPSS) in the near integration using (quasi) GLS-detrending[1]. Asymptotic properties of the test in the constant case were compared with point-optimal test proposed by Müller (2005) for different initial conditions. The results revealed that in the case of small initial conditions the test proposed in Müller (2005) is effective. Given a large, or even moderate initial conditions, this test has serious liberal size distortions, tending to unity for strongly autocorrelated stationary data generating processes. Additionally, it is dominated by the HLM test with large initial conditions.

In this paper, we consider the asymptotic properties of stationarity tests proposed by HLM and Müller (2005) following the approach of HLT. In Section 2 we introduce the HLM test and point-optimal test proposed in Müller (2005) and obtain corresponding limiting distributions assuming local behavior of the trend. Additionally, we use parametrization of the initial condition following the Müller and Elliott (2003) approach. In Section 3.1 we analyze these tests in the case of an asymptotically negligible initial condition, assuming the local behavior of the trend. As in HLM, in this case the asymptotic size curves show that the point-optimal tests of Müller (2005) is superior to the HLM test. Additionally, the test with only constant has serious liberal size distortions, driven by increase of the magnitude of the local trend parameter. We propose to use the intersection of rejections testing strategy [2] of two tests with and without a trend in deterministic component, i.e., we reject the null hypothesis, if both tests simultaneously reject it. We also propose a modification of this decision rule, using pre-testing of a trend parameter and using this information to perform only the test with trend, if there is evidence of a large local trend. As simulations show, this procedure has advantages over a simple intersection rejection of two test (with- and without-trend). In Section 3.2 we analyze a similar procedure, assuming knowledge of deterministic term, but no knowledge of initial condition magnitude. In this case, as in Section 3.1, simple intersection of rejections of corresponding tests is the best solution, as well as the modification with pre-testing of the initial condition. In Section 3.3 we address the composite problem of uncertainty over both the linear trend and initial condition. Following HLT, we propose an intersection of rejections testing strategy consisting of all four test statistics, as well as modification in the pre-testing the trend magnitude and initial condition. Asymptotic analysis reveals

[1]Müller (2005) revealed that the use of conventional KPSS test with the bandwidth parameter in the long-run variance estimator increasing at a slower rate than the length of the sample, leads to an asymptotic size equal to unity under the null hypothesis about near integration. Therefore, in our analysis (local to unit root) we do not consider conventional OLS-detrended KPSS test.

[2]HLT used the "union of rejections" term, and their test rejected the null hypothesis, if *at least one of the tests* rejected it. As we consider stationarity tests in our procedure, we reject the null hypothesis if *all of tests* reject it, and we call this strategy the "intersection of rejections".

considered modifications as superior to the use of the individual tests with varying parameters in the local trend and the initial condition.

In our analysis, the size of all tests was compared to a fixed power, where critical values were obtained for the integrated process, as well as with zero trend parameters and zero initial conditions. In Section 4 we obtain the critical values and scaling constants for fixed amount of mean reversion under the stationary null hypothesis and discuss the behavior of tests in finite samples (finite sample results are available in the online Appendix to this paper). Results obtained in the course of this investigation are formulated in the Conclusion section.

2 The Model

We consider the data generating process (DGP) as

$$y_t = \mu + \beta t + u_t,\ t = 1, \ldots, T, \quad (1)$$
$$u_t = \rho u_{t-1} + \varepsilon_t,\ t = 2, \ldots, T, \quad (2)$$

where the process ε_t is taken to satisfy the standard assumptions considered by Phillips and Solo (1992):

Assumption 1 *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* $\sup_t \mathbb{E}(e_t^4) < \infty$*. Short-run and long-run variances 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 contrast to the conventional KPSS test, we use the local to unity behavior of the parameter ρ, i.e., $\rho = \rho_T = 1 - c/T$, where $c \geq 0$. We consider the local asymptotics primarily because it provides a more accurate approximation for small samples than the standard asymptotics when the series contains a large autoregressive root frequently observed in many macroeconomic time series. We test the null of stationarity (local to unit root) $H_0 : c \geq \bar{c} > 0$ against alternative hypothesis $H_1 : c = 0$, where $\bar{c}$ is the minimal amount of mean reversion under the stationary null hypothesis.

Conventional KPSS test with the bandwidth parameter in the long-run variance estimator increasing at a slower rate than the length of the sample leads to an asymptotic size equal to unity under the null hypothesis about near integration (see Müller (2005)). Thus, we consider two tests having non-degenerate limiting distribution under local to unit root behavior of autoregressive parameter. The first was suggested by Müller (2005). Following Müller and Elliott (2003), Müller (2005) proposed an asymptotically optimal test statistic $Q^\mu(\bar{c})$ for the constant case and $Q^\tau(\bar{c})$ for the trend case to discriminate between a null hypothesis $\rho_T = 1 - c/T$ and $\rho_T = 1$. This statistic is constructed as:

$$Q^i(\bar{c}) = q_1^i(\hat{\omega}_\varepsilon^{-1}T^{-1/2}\hat{u}_T^i)^2 + q_2^i(\hat{\omega}_\varepsilon^{-1}T^{-1/2}\hat{u}_1^i)^2 + q_3^i(\hat{\omega}_\varepsilon^{-1}T^{-1/2}\hat{u}_T^i)(\hat{\omega}_\varepsilon^{-1}T^{-1/2}\hat{u}_1^i) + q_4^i\hat{\omega}_\varepsilon^{-2}T^{-2}\sum_{t=1}^{T}(\hat{u}_t^i)^2, \quad (3)$$

where $\hat{u}_t^i$ are OLS residuals from regression of y_t on d_t, where $d_t = \mu$ in the constant case and $d_t = \mu + \beta t$ in the trend case, $q_1^\mu = q_2^\mu = \bar{c}(1+\bar{c})/(2+\bar{c})$, $q_3^\mu = -2\bar{c}/(2+\bar{c})$, $q_4^\mu = \bar{c}^2$ and $q_1^\tau = q_2^\tau = \bar{c}^2(8+5\bar{c}+\bar{c}^2)/(24+24\bar{c}+8\bar{c}^2+\bar{c}^3)$, $q_3^\tau = 2\bar{c}^2(4+\bar{c})/(24+24\bar{c}+8\bar{c}^2+\bar{c}^3)$, $q_4^\tau = \bar{c}^2$. Additionally, $\hat{\omega}_\varepsilon^2$ is any consistent estimator of long-run variance of ε_t using residuals from an $AR(1)$ regression of $\hat{u}_t^i$.[3]

The second test was proposed by HLM. It uses the (quasi) GLS-detrended series. Specifically, $\tilde{u}_t^i, i = \mu, \tau$ are OLS residuals from regression of $y_{\bar{c}} = y_t - \bar{\rho}_T y_{t-1}$ on $Z_{\bar{c}} = z_t - \bar{\rho}_T z_{t-1}$, $t = 2, \ldots, T$, where $z_t = 1$ in the constant case and $z_t = (1, t)'$ in the trend case. Then, the $S^i(\bar{c})$ test statistic is constructed as

$$S^i(\bar{c}) = \frac{T^{-2}\sum_{t=2}^{T}\left(\sum_{j=2}^{t}\tilde{u}_j^i\right)^2}{\hat{\omega}_\varepsilon^2}, \tag{4}$$

where kernel-based long-run variance estimator $\hat{\omega}_\varepsilon^2$ is calculated using GLS-detrended residuals $\tilde{u}_t^i$.[4]

We consider the following two assumption, specifying the behavior of trend coefficient β and initial condition u_1.

Assumption 2 *The trend coefficient β satisfies $\beta = \beta_T = \kappa\omega_\varepsilon T^{-1/2}$, where κ is some finite constant.*

Assumption 3 *The initial condition u_1 satisfies $u_1 = \xi = \alpha\sqrt{\omega_\varepsilon^2/(1-\rho_T^2)}$, where $\rho_T = 1 - c/T$, $c > 0$. In unit root case, $c = 0$, the initial condition is equal to zero, i.e. all tests are invariant to α.*

HLM showed that in the case of small initial conditions the $Q^i(\bar{c})$ test is efficient, in terms of its smaller size, in comparison with the $S^i(\bar{c})$ test. However, for large initial conditions $Q^i(\bar{c})$ has serious size distortion and is strictly dominated by $S^i(\bar{c})$ test. One could also consider the optimal $Q^i(g,k)$ test proposed by Elliott and Müller (2006) and analyzed by HLM in a stationarity testing context. However, as it was shown in HLM, this test is strictly dominated by $Q^i(\bar{c})$ test for small initial conditions and by $S^i(\bar{c})$ test for large initial conditions. Thus, we do not use it in further consideration. Similarly, as will be shown in Section 3.1, in the absence of trend, effective tests are those that do not account for the trend in construction, while these tests will have a serious size distortion in the presence of a trend in DGP.

The following lemma summarizes limiting distributions of four tests under Assumptions 1-3.

Lemma 1 *Let $\{y_t\}$ be generated as in* (1) *and* (3) *and Assumptions 1-3 be satisfied. Then under $\rho_T = 1 - c/T$, $0 \le c < \infty$*

$$\begin{aligned} Q^\mu(\bar{c}) &\Rightarrow q_1^\mu\left(K_c^\mu(1)+\frac{\kappa}{2}\right)^2 + q_2^\mu\left(K_c^\mu(0)-\frac{\kappa}{2}\right)^2 \\ &+ q_3^\mu\left(K_c^\mu(1)+\frac{\kappa}{2}\right)\left(K_c^\mu(0)-\frac{\kappa}{2}\right) + q_4^\mu\int_0^1\left(K_c^\mu(r)+\kappa(r-\frac{1}{2})\right)^2 dr, \end{aligned} \tag{5}$$

$$Q^\tau(\bar{c}) \Rightarrow q_1^\tau K_c^\tau(1)^2 + q_2^\tau K_c^\tau(0)^2 + q_3^\tau K_c^\tau(1)K_c^\tau(0) + q_4^\mu\int_0^1 K_c^\tau(r)dr, \tag{6}$$

[3] Kernel-based long-run variance estimator proposed by (Müller, 2005) and used by HLM leads to very poor finite sample properties of the tests under negative moving average errors. In unreported simulations (results made available upon request) we found that autoregressive long-run variance estimator leads to satisfactory finite sample properties.

[4] Note that this estimator has satisfactory finite sample properties in unreported finite sample simulations.

$$
\begin{aligned}
S^{\mu}(\bar{c}) &\Rightarrow \int_0^1 H_{c,\bar{c},\alpha,\kappa}(r)^2 dr, &(7)\\
S^{\tau}(\bar{c}) &\Rightarrow \int_0^1 \left(H_{c,\bar{c},\alpha,0}(r) - 6r(1-r)\int_0^1 H_{c,\bar{c},\alpha,0}(s)ds\right)^2 dr, &(8)
\end{aligned}
$$

Here

$$
\begin{aligned}
K_c^{\mu}(r) &= K_c(r) - \int_0^1 K_c(s)ds,\\
K_c^{\tau}(r) &= K_c^{\mu}(r) - 12\left(r - \frac{1}{2}\right)\int_0^1 \left(s - \frac{1}{2}\right)K_c(s)ds,\\
H_{c,\bar{c},\alpha,\kappa}(r) &= K_c(r) + \bar{c}\int_0^r (K_c(s) + \kappa s)ds - r\left[K_c(1) + \bar{c}\int_0^1 (K_c(s) + \kappa s)ds\right],\\
K_c(r) &= \begin{cases}\alpha(e^{-rc} - 1)(2c)^{-1/2} + W_c(r), & c > 0\\ W(r), & c = 0\end{cases},
\end{aligned}
$$

where $W_c(r) = \int_0^r e^{-(r-s)c}dW(s)$ *is a Ornstein-Uhlenbeck process,* $W(r)$ *is a standard Wiener process, and* $\Rightarrow$ *denotes weak convergence.*

The proof of (5) is similar to Harvey *et al.* (2009), the proof of (7) is given in the Appendix. Proofs of (6) and (8) are standard and use FCLT and CMT. Also, following HLM (see also Müller and Elliott (2003) and Elliott and Müller (2006)), we set $\bar{c} = 10$ for $Q^{\mu}(\bar{c})$ and $S^{\mu}(\bar{c})$ statistics and $\bar{c} = 15$ for $Q^{\tau}(\bar{c})$ and $S^{\tau}(\bar{c})$ statistics.

Remark 1 Note that the $Q^{\tau}(\bar{c})$ and $S^{\tau}(\bar{c})$ are invariant to the trend magnitude while the limiting distributions of $Q^{\mu}(\bar{c})$ and $S^{\mu}(\bar{c})$ explicitly depend on the local drift parameter κ.

Remark 2 Under a fixed nonzero trend of the form $\beta = \kappa\omega_{\varepsilon} \neq 0$ it is easy to show that the $Q^{\mu}(\bar{c})$ and $S^{\mu}(\bar{c})$ are both $O_p(T)$, i.e. diverge to infinity. This leads to the fact that these tests have size equal to unity for all c.

3 Asymptotic analysis of stationarity tests

3.1 Asymptotic behavior under a local trend

Consider a case when the initial condition is $u_1 = o_p(T^{1/2})$. Then, in limiting distributions obtained in Lemma 1 the process $K_c(r)$ is simply replaced by Ornstein-Uhlenbeck process $W_c(r)$. Figures 1(a)-(d) show asymptotic size for $c \in [0, 20]$, where the critical conditions are obtained for test comparison for $c = 0$ and $\kappa = 0$ in order for the power to equal 0.5 for any test as in Müller (2005) and HLM[5].

Comparing the size of tests for the case of $\kappa = 0$ (fig. 1(a)), under the absence of a linear trend, it is clear that the size is smaller for tests that do not take into account the presence of a trend. Additionally, the Q^i dominates S^i, as it was evident in the results of simulations performed

[5]Here and in the following sections results are obtained by simulations of the limiting distributions in Lemma 1, 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 50,000 replications.

by Harris *et al.* (2007) for the case of the only constant in deterministic term. As a result, for stationarity testing, the Q^μ test will be most effective of the tests considered for $\kappa = 0$.

For $\kappa = 0.5$ (fig. 1(b)) the results are largely similar to the case of $\kappa = 0$ and Q^μ is still effective, except in a very small interval $c \in [0, 1]$, when Q^μ is dominated by tests that take into account the trend (though they have a higher power in the interval $c \in [0, 1]$ and this interval can be considered as negligible). Notably, that size of tests with only the constant slightly increases in comparison with case of $\kappa = 0$. For $\kappa = 1$ (fig. 1(c)) the S^τ and Q^τ are clearly superior to the S^μ and Q^μ (the latter have serious size distortions, the size is never lower than 0.4 for the considered interval $c \in [0, 20]$), and Q^τ is the efficient test. The size of tests with only the constant continues to increase with increasing κ. For $\kappa = 2$ (fig. 1(d)) the size of S^μ and Q^μ almost always equals to one, which confirms the behavior Q^μ and S^μ under a fixed nonzero trend (see Remark 2).

As each of the considered Q^μ and Q^τ tests is efficient (in terms of size) for some values of a local trend, it is necessary to use a feasible strategy to discriminate between the two cases (presence/absence of a linear trend), if there is uncertainty over the magnitude of this local trend.

Following Harvey *et al.* (2009) we reject the null of stationarity if both tests reject the null simultaneously. Specifically, this intersection of rejections decision rule can be written as:

$$IR = \text{Reject } H_0 \text{ if } \{Q^\mu > m_\xi cv_\xi^{Q,\mu} \text{ and } Q^\tau > m_\xi cv_\xi^{Q,\tau}\}, \tag{9}$$

where $cv_\xi^{Q,\mu}$ and $cv_\xi^{Q,\tau}$ are the asymptotic critical values for Q^μ and Q^τ for some specified value of c and significance level ξ (for more details see Section 4), and m_ξ is some scaling constant ensuring that asymptotic size equals ξ for a given value c (in case of absence of scaling the size and power decreases, so we call the decision rule with scaling liberal).

It is possible to further improve this strategy by using information about the large value of parameter κ, specifically, about the clear evidence of a trend. As noted by J. Breitung in rejoinder of Harvey *et al.* (2009) there is no need to use the same scaling constant m_ξ in all cases. If there is strong evidence for a trend, then the probability of rejection of Q^μ test tends to unity. Consequently, it is possible to improve the strategy IR by using pre-tests $Dan\text{-}J$, t_λ, t_λ^{m2} and t_β^{RQF},[6] proposed, respectively, by Bunzel and Vogelsang (2005), Harvey *et al.* (2007), Perron and Yabu (2009) and analyzed in HLT and Harvey *et al.* (2010) for various magnitudes of a local trend κ and initial condition α (i.e. with strong evidence for the trend, the critical value of the Q^τ can be used). Specifically, consider the following modification of the decision rule (9):

$$IR(s_\beta) = \begin{cases} \text{Reject } H_0 \text{ if } \{Q^\mu > m_\xi cv_\xi^{Q,\mu} \text{ and } Q^\tau > m_\xi cv_\xi^{Q,\tau}\}, & \text{if } |s_\beta| \le cv_\beta \\ \text{Reject } H_0 \text{ if } \{Q^\tau > cv_\xi^{Q,\tau}\}, & \text{if } |s_\beta| > cv_\beta \end{cases}, \tag{10}$$

where s_β denotes some pre-test statistic for testing $\beta = 0$, and cv_β is the corresponding critical value. The limiting distribution of these two tests follows directly from Lemma 1 and CMT, and is, therefore, omitted. Under a fixed nonzero trend each of s_β statistics diverges to infinity, so that asymptotically the Q^τ test will be selected. On the other hand, in the absence of a trend the Q^μ with scaling critical value will actually be selected.

Figures 2(a)-(d) show asymptotic size of tests Q^μ, Q^τ, IR, $IR(|t_\lambda|)$, $IR(|t_\lambda^{m2}|)$ and $IR(|Dan\text{-}J|)$ for values $\kappa \in \{0, 1, 2, 4\}$ and $c \in [0, 20]$. We chose the scaling constant m_ξ in order for the asymptotic power of the test IR to equal 0.50 for $\kappa = 0$. For $\kappa = 0$ the size of IR tests is between Q^μ

[6] The t_λ and t_β^{RQF} statistics are asymptotic equivalents, therefore we consider only the first in a study of the asymptotic behavior of the tests.

and Q^τ, as expected. The size of $IR(s_\beta)$ tests is almost identical to IR. Size curves of $IR(s_\beta)$ nearly coincide with IR, as the hypothesis of $\beta = 0$ has rejected very rarely. For $\kappa = 1$, rejection of the hypothesis of $\beta = 0$ is still quite rare, so that the size curves of $IR(s_\beta)$ are close toIR, and size curve of IR lies between ones of Q^μ and Q^τ. With increasing κ, the size of tests $IR(s_\beta)$ approaches Q^τ, the effective test in this case. The size curve $IR(t_\lambda)$ for $\kappa = 4$ coincides with Q^τ, as even for $\kappa \approx 2.6$ the hypothesis of $\beta = 0$ will almost always be rejected by the t_λ (see Harvey *et al.* (2008, Fig. 5(a) and 6(a))). Comparing $IR(t_\lambda^{m2})$ and $IR(Dan\text{-}J)$, the size curve of the latter is closer to Q^τ, as for large κ the $Dan\text{-}J$ test rejects the hypothesis of $\beta = 0$ less frequently than t_λ^{m2} (see Harvey *et al.* (2008, Fig. 5(a) and 6(a))).

3.2 Asymptotic behavior under various initial conditions

We consider tests Q^i and S^i, $i = \mu, \tau$, in an approach similar to the previous section, by varying parameters of the initial condition α from 0 to 6 [7] and varying parameter $c \in \{5, 10, 15\}$. Additionally, we assume knowledge of the deterministic component type. Figures 3(a) and (c) show asymptotic size of the tests Q^μ, S^μ, IR and $IR(s_\alpha)$. The last two tests will be discussed below. Our results show that for small initial conditions the Q^μ test dominates S^μ, while with increasing α the asymptotic size of Q^μ test goes to one (for moderate values of c). At the same time, the size of S^μ increases with increasing α only for small values of c, but for $c = 10$ (Fig. 3(c)) it remains constant for all α, even though it is dominated by Q^μ for small initial conditions ($\alpha < 2.6$). Thus, for small α the Q^μ test is efficient, while for large values of α it is strongly oversized. Therefore, if information is available about the large initial condition, it becomes necessary to use the S^μ test.

Our results for the trend case are summarized in Figures 3(b) and (d) for $c = 10$ and $c = 15$, respectively, and appear to be identical, although the size distortions for Q^τ and S^τ are not as strong for large α and small c, as in Q^μ and S^μ (results available on request).

As in Harvey *et al.* (2009) the following strategy of intersection of rejection can be applied:

$$IR = \text{Reject } H_0 \text{ if } \{Q^i > m_\xi^i cv_\xi^{Q,i} \text{ and } S^i > m_\xi^i cv_\xi^{S,i}\}, \tag{11}$$

where $cv_\xi^{Q,i}$ and $cv_\xi^{S,i}$, $i = \mu, \tau$ are the asymptotic critical values of tests Q^i and S^i for some specified value c and significance level ξ, and m_ξ^i is the some scaling constant in order for the asymptotic size to be at ξ level for a given c.

In an approach that is similar to the previous section, this strategy can be modified by using additional information about the large initial condition, in order to select only the test S^i in this case:

$$IR(s_\alpha) = \begin{cases} \text{Reject } H_0 \text{ if } \{Q^i > m_\xi^i cv_\xi^{Q,i} \text{ and } S^i > m_\xi^i cv_\xi^{S,i}\}, & \text{if } s_\alpha \le cv_\alpha \\ \text{Reject } H_0 \text{ if } \{S^i > cv_\xi^{S,i}\}, & \text{if } s_\alpha > cv_\alpha \end{cases}, \tag{12}$$

where s_α denotes some test statistic for testing $\alpha = 0$, and cv_α is the corresponding critical value. As in the previous section, the limiting distribution of these two tests directly follows from Lemma 1 and CMT and omitted for brevity.

It can be used the following statistic (proposed by HLT) for s_α:

$$s_\alpha = DF\text{-}QD^i - \frac{cv_\xi^{QD,i}}{cv_\xi^{OLS,i}} DF\text{-}OLS^i, \tag{13}$$

[7] As these tests are symmetric around α there is no need to consider negative initial conditions.

where $i = \mu, \tau$ depends on the type of the deterministic term, $DF\text{-}QD^\tau$ and $DF\text{-}OLS^\tau$ are ADF tests for GLS and OLS detrended data, respectively, and $cv_\xi^{QD,\tau}$ and $cv_\xi^{OLS,\tau}$ are corresponding critical values. Large values of the upper tail of distribution for this test statistic indicate a large initial condition of $|\alpha|$. HLT obtained the critical values for $c = 30$; only in this case the test statistic s_α had the correct size. For smaller values of c it contained liberal size distortions that increased with decreasing c.

Figures 3(a)-(d) also show the asymptotic size of IR and $IR(s_\alpha)$ tests. As in the previous section the critical values of Q^i and S^i ($i = \mu, \tau$) and scaling factors m'_ξ were calculated in such a way, that the tests had their powers equal to 0.50. We use the critical values for s_α at $c = 10$ for the mean case and at $c = 20$ for the trend case (these values provide better test properties). We also analyzed the behavior of the tests, if the critical values for s_α were obtained at different c (results are available upon request). Furthermore, we use an additional correction of critical values, so that the $IR(s_\alpha)$ strategy has power equal to 0.50. As expected, the size curve of IR lies between the size curves of S^i and Q^i, although the size distortion is quite substantial for large α. However, the modification $IR(s_\alpha)$ has a size that is close enough to effective S^i test for large α and significant gain in size for small α in comparison to S^i. Thus, $IR(s_\alpha)$ effectively discriminates the cases of small and large initial conditions.

3.3 Asymptotic behavior under uncertainty over both the trend and initial condition

Previous sections of this work discussed two testing strategies, the first of which is a stationarity test with uncertainty over linear trend with the knowledge about asymptotically negligible initial condition. The second testing strategy is stationarity test with the knowledge of the exact specifications of the deterministic component, but with uncertainty over the magnitude of the initial condition. The magnitude of the initial condition or the value of a trend parameter are not known in advance, however. In this case following the HLT method (see also Harvey *et al.* (2008)), we can apply the strategy of intersection of rejections, which reject stationarity, if each of the four tests, Q^i and S^i ($i = \mu, \tau$), rejects the null hypothesis of stationarity. This liberal decision rule can be written as:

$$IR_4 = \text{Reject } H_0 \text{ if } \{Q^\mu > m_\xi^* cv_\xi^{Q,\mu} \text{ and } Q^\tau > m_\xi^* cv_\xi^{Q,\tau} \text{ and } S^\mu > m_\xi^* cv_\xi^{S,\mu} \text{ and } S^\tau > m_\xi^* cv_\xi^{S,\tau}\}, \quad (14)$$

where m_ξ^* is the scaling constant. It can also improve the test pre-identifying the possible large initial condition or significant linear trend as in Sections 3.1 and 3.2. However, as shown in Harvey *et al.* (2008), the tests $Dan\text{-}J$, t_λ and t_λ^{m2} are very sensitive to the magnitude of initial condition. HLT used a modified test t'_λ:

$$t'_\lambda = (1 - \lambda^*) t_0 + \lambda^* t_1, \quad (15)$$

where in contrast to Harvey *et al.* (2007), t_1 is constructed using autocorrelation-corrected t-ratio for testing $\beta_T = 0$ in the regression

$$y_t - \bar{\rho}_T y_{t-1} = \hat{\mu}(1 - \bar{\rho}_T) + \hat{\beta}_T(t - \bar{\rho}_T(t-1)) + \hat{u}_t,\ t = 2, \ldots, T, \quad (16)$$

where $\bar{\rho}_T = 1 - \bar{c}/T$, and corresponding long-run variance is calculated by using residuals $\hat{u}_t$. HLT obtained the limiting distribution of the modified test statistic and showed that for $c = \bar{c}$

this statistic is asymptotically invariant to the initial condition at point $c = \bar{c}$ and asymptotically standard normal. HLT set $s_\beta = |t'_\lambda|$ with $\bar{c} = 30$ and used the standard normal critical value. Then the test s_β will be oversized as c decreases towards zero, but at $c = 30$ will be correctly sized.

Thus, as in HLT, the modified liberal decision rule can be written as follows, where each $IR(\cdot, \cdot)$ strategy reject the null hypothesis if all tests in the brackets reject it:

Definition 1 *The modified intersection of rejections strategy IR_4^* is defined as follows:*

1) *If $s_\beta \le cv_\beta$ and $s_\alpha \le cv_\alpha$, then use the liberal decision rule $IR(Q^\mu, Q^\tau, S^\mu, S^\tau)$, defined in* (14)*;*

2) *If $s_\beta \le cv_\beta$ and $s_\alpha > cv_\alpha$, then use the liberal decision rule $IR(S^\mu, S^\tau)$, the corresponding scaling constant is defined as m_ξ^{**};*

3) *If $s_\beta > cv_\beta$ and $s_\alpha \le cv_\alpha$, then use the liberal decision rule $IR(Q^\tau, S^\tau)$, the corresponding scaling constant is defined as m_ξ^τ;*

4) *If $s_\beta > cv_\beta$ and $s_\alpha > cv_\alpha$, then use the decision rule reject H_0, if $S^\tau > cv_\xi^{S,\tau}$.*

The basic concepts of this strategy are as follows. Under 1) there is no reason to assume that the values of the local trend and the initial condition are large, and there is no reason to argue that they are necessarily small. Thus, the IR_4 strategy should be applied. Under 2) there may be some evidence in favor of a large initial condition, so that Q^μ and Q^τ tests will be ineffective (then can have size approaching unity) and can be excluded from the intersection of rejections IR_4. However, since we can not be sure of the magnitude of the local trend, it is necessary to use both tests S^μ and S^τ. Under 3) there is evidence of a large magnitude of a local trend. In this case, Q^μ and S^μ are ineffective and should be excluded from the IR_4 strategy. However, we should consider both Q^τ and S^τ tests, as there is no reason to believe that the initial condition is large. Finally, under 4) there is evidence of both a large local trend and large initial condition so that only S^τ will be effective test in this case. Thus, the null hypothesis will be rejected if only this test will be reject it.

Figures 4-6 respectively demonstrate for $\kappa = 0, 0.5, 1$, the asymptotic size of tests S^τ, IR_4 and IR_4^* for $c \in \{0, 1, \ldots, 20\}$ by fixing power at 0.50. Results for larger κ are discussed below. All figures (a)-(i) show the results for $\alpha = \{-4, -2, -1, -0.5, 0, 0.5, 1, 2, 4\}$, respectively. Only S^τ is considered in the figures among the four original stationarity tests, as its size is never below a certain level across considered α and κ. Thus, this test can be considered "robust" under both forms of uncertainty (trend and initial condition). Notably, it is necessary to correct the testing strategy IR_4^*, so that it controls size asymptotically as in HLT. Therefore, we replace $cv_\xi^{Q,\mu}$, $cv_\xi^{Q,\tau}$, $cv_\xi^{S,\mu}$ and $cv_\xi^{S,\tau}$ by $\tau_\xi cv_\xi^{Q,\mu}$, $\tau_\xi cv_\xi^{Q,\tau}$, $\tau_\xi cv_\xi^{S,\mu}$ and $\tau_\xi cv_\xi^{S,\tau}$, respectively, where τ_ξ is the scaling constant, such that the power has never been lower than 0.50.

When $\kappa = 0$ the IR_4^* test outperforms the S^τ test. Only in case of $\alpha = 4$ for small c the size of IR_4^* is slightly higher. In comparison with the IR_4 test, its size is almost always smaller than the S^τ and IR_4^* tests, although in case of a large $|\alpha|$, this test has serious size distortions for small c. When $\kappa = 0.5$ for negative α the size of IR_4^* is slightly higher than S^τ (except for $c > 5$ in case of $\alpha = -4$), but increasing α, even at $\alpha = 0.5$ causes their size curves to intersect, and for $\alpha = 2$ the size of IR_4^* is lower than S^τ. For $|\alpha| < 1$ the size of IR_4 behaves almost as well as S^τ, but for large $|\alpha|$ its size curves become strongly nonmonotonic.

In our discussion of results of larger κ, it should be noted that when $\kappa = 0$, IR_4 outperforms two of the other tests and S^τ shows serious size distortion. When $\kappa = 1$, the S^τ becomes dominant, while IR_4 shows considerable size distortion. In both cases the size curve of IR_4^* lies between the ones of S^τ and IR_4. According to Definition 1, as κ becomes larger, S^τ will continue to dominate, and the size curve of IR_4^* will lie between the ons of S^τ and IR_4, but will get close to S^τ (however, due to scaling, it will not coincide with S^τ even for very large κ). Therefore, under uncertainty about the trend, IR_4^* testing strategy should be applied, whose size distortion is acceptable regardless of whether a large time trend exists.

Thus, based on the asymptotic results, we strongly recommend the use of the modified decision rule IR_4^*, if there is uncertainty over both the trend and the initial condition.

4 Critical values

In this section we discuss obtaining critical values for their practical application[8], as previously we compared the tests by fixing the power at 0.50.

The critical values for tests $Q^\mu(\bar{c})$ and $S^\mu(\bar{c})$ are given in Table 1. We obtain them at $c = 10$ for tests $Q^\mu(\bar{c})$ and $S^\mu(\bar{c})$, as in Müller and Elliott (2003) and Elliott and Müller (2006), and for tests $Q^\tau(\bar{c})$ and $S^\tau(\bar{c})$ at $c = 15$. We note (see HLM), that in this case $c = \bar{c}$, and tests $S^i(\bar{c})$, $(i = \mu, \tau)$ have standard KPSS limiting distributions, so the critical values are the same as for conventional KPSS tests. However, the finite sample critical values slowly converge to the asymptotic, so we provide additional critical values of all tests for $T = 150, 300, 600$. Asymptotic scaling constants, however, are appropriate for finite samples. Also, $Q^i(\bar{c})$ is not invariant to the initial condition α, when $c > 0$, so the critical values obtained for $\alpha = 0$.

Critical values for the initial condition test s_α were obtained at $c = 20$ and are listed in Table 2.

It is necessary to obtain scaling constants for all intersection of rejections testing strategy, which were considered in Sections 3. However, there is some difficulty with obtaining the scaling constants, as the critical values for the tests with trend are constructed with $c = 15$, while critical values for the tests only with constant, are constructed with $c = 10$. Therefore, if the intersection of rejections testing strategy includes tests with different types of deterministic components, we obtain the scaling constants with $c = 12.5$ (i.e. at $c = 12.5$ tests has correct size). Otherwise, scaling is performed using the specific c for each type of deterministic component. All scaling constants are given in Table 3 (the simulation code is available upon request).

We also conduct finite sample simulations for proposed strategies (by fixing the power at 0.50 to allow comparison), by using QS kernel and automatic bandwidth selection of Newey and West (1994) for long-run variance estimator of $S^i(\bar{c})$ statistics and autoregressive long-run variance estimator of $Q^i(\bar{c})$ statistics. Various DGP were examined ($i.i.d.$, $AR(1)$ and $MA(1)$ processes for ε_t). Results show that asymptotic analysis provides a good approximation for the behavior of tests in finite samples (results available in the online Appendix to this paper). Notably, only $S^i(\bar{c})$ and $Q^i(\bar{c})$ requires finite sample critical values while the asymptotic scaling constants are similar to the asymptotic.

[8] In this section, we obtain the results using the normalized sum of 5,000 steps and 100,000 replications.

5 Conclusion

In this paper we considered the problem of stationarity testing with uncertainty over the trend and/or initial condition. We proposed the intersection of rejections testing strategy of several tests (similar to using a union of rejection testing strategy proposed by HLT), if there is uncertainty over the trend and/or initial condition. Simulation evidence revealed that a testing strategy based on the rejection of all tests each of which is effective for small/large initial conditions and/or for small/large parameter of the local trend suggests size performance in the presence of uncertainty over both the trend and the initial condition. Additionally, we showed that pre-testing of the trend parameter and of the initial condition could improve the procedure if any of these parameters will be significantly nonzero. As stationarity testing is necessary for confirmatory analysis, we conclude that our procedure is useful in empirical applications in conjunction with the HLT test.

Appendix

Proof of Lemma 1: Due to the invariance, we set $\mu = 0$ without loss of generality, thus $y_t = \beta t + u_t$.

We consider a model without the trend, when it is actually present. Using GLS-detrending we obtain that

$$\begin{aligned} r_t &= y_t - \bar{\rho} y_{t-1} - \tilde{\mu}, \\ \tilde{\mu} &= \frac{1}{T-1}\sum_{t=2}^{T}(y_t - \bar{\rho} y_{t-1}). \end{aligned}$$

Let $z_t = u_t - u_1$. Then

$$\begin{aligned} r_t &= u_t - \bar{\rho} u_{t-1} + \beta t - \bar{\rho}\beta(t-1) - \frac{1}{T-1}\sum_{t=2}^{T}(u_t - \bar{\rho} u_{t-1}) - \frac{\beta}{T-1}\sum_{t=2}^{T}(t - \bar{\rho} t + \bar{\rho}) \\ &= z_t + u_1 - \bar{\rho} z_{t-1} - \bar{\rho} u_1 - \frac{1}{T-1}\sum_{t=2}^{T}(z_t + u_1 - \bar{\rho} z_{t-1} - \bar{\rho} u_1) \\ &+ \beta t - \bar{\rho}\beta(t-1) - \frac{\beta}{T-1}\sum_{t=2}^{T}(t - \bar{\rho} t + \bar{\rho}) \\ &= \left[z_t - \bar{\rho} z_{t-1} - \frac{1}{T-1}\sum_{t=2}^{T}(z_t - \bar{\rho} z_{t-1}) \right] + \left[\beta t - \bar{\rho}\beta t - \frac{\beta}{T-1}\sum_{t=2}^{T}(t - \bar{\rho} t) \right] \end{aligned}$$

The second term can be written as $\beta\left(t - \frac{T+2}{2}\right) - \bar{\rho}\beta\left(t - \frac{T+2}{2}\right) = T^{-1}\bar{c}\beta\left(t - \frac{T+2}{2}\right)$.

Then

$$\begin{aligned} T^{-1/2}\sum_{i=2}^{[rT]} r_t = &\left[T^{-1/2}\sum_{i=2}^{[rT]}\left(z_t - \bar{\rho} z_{t-1} - \frac{1}{T-1}\sum_{t=2}^{T}(z_t - \bar{\rho} z_{t-1}) \right) \right] \\ &+ \left[T^{-1/2}\sum_{i=2}^{[rT]} T^{-1}\bar{c}\beta\left(i - \frac{T+2}{2} \right) \right] \end{aligned}$$

The first term of this expression has a limiting distribution obtained in HLM and corresponds to the case of $\kappa = 0$ (more precisely, converges to $\omega_\varepsilon H_{c,\bar{c},\alpha,0}(r)$). Consider the second term, responsible for the behavior of test statistics under local trend:

$$T^{-1/2}\sum_{i=2}^{[rT]} T^{-1}\bar{c}\beta\left(i - \frac{T+2}{2}\right) = \bar{c}\omega_\varepsilon\kappa T^{-2}\frac{([rT]-1)([rT]+2)}{2} - \bar{c}\omega_\varepsilon\kappa T^{-2}\frac{([rT]-1)(T+1)}{2}$$

This expression converges to $\bar{c}\omega_\varepsilon\kappa\left(\frac{r^2}{2} - \frac{r}{2}\right)$, which proves the lemma using CMT and simple integral transformations, because the long-run variance estimators (kernel-based in our case) ω_ε are still consistent under the local trend misspecification.

■

References

Bunzel, H. and Vogelsang, T.J. (2005). Powerful Trend Function Tests That Are Robust to Strong Serial Correlation with an Application to the Prebisch-Singer Hypothesis. *Journal of Business and Economic Statistics*, **23**, 381–394.

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

Harris, D., Leybourne, S., and McCabe, B. (2007). Modified KPSS Tests for Near Integration. *Econometric Theory*, **23**, 355–363.

Harvey, D.I., Leybourne, S.J., and Taylor, A.M.R. (2007). A Simple, Robust and Powerful Test of the Trend Hypothesis. *Journal of Econometrics*, **141**, 1302–1330.

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). 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). Testing for Unit Roots in the Presence of Uncertainty over Both the Trend and Initial Condition. *Journal of Econometrics*, **169**, 188–195.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root? *Journal of Econometrics*, **54**, 159–178.

Maddala, G.S. and Kim, I-M (1998). *Unit Roots, Cointegration, and Structural Change*. Cambridge University Press, Cambridge.

Müller, U. (2005). Size and Power of Tests for Stationarity in Highly Autocorrelated Time Series. *Journal of Econometrics*, **128**, 195–213.

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

Newey, W.K. and West, K.D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, **61**, 631–653.

Perron, P. and Yabu, T. (2009). Estimating Deterministic Trends with an Integrated or Stationary Noise Component. *Journal of Econometrics*, **151**, 56–69.

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

Table 1. Asymptotic and finite sample critical values for Q^i and S^i, $i = \mu, \tau$ at the ξ significance level

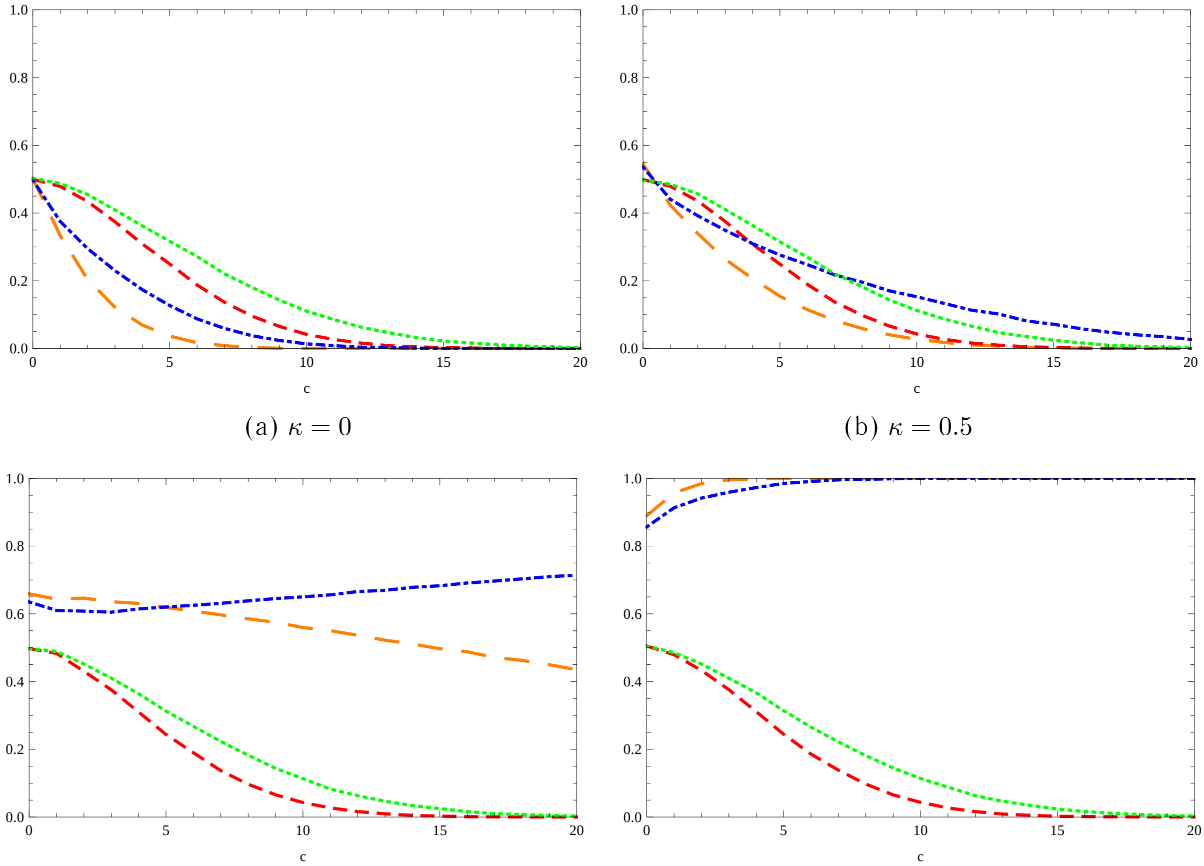
	T	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$
Q^μ	∞	6.93	8.04	10.55
	600	6.90	8.07	10.65
	300	6.98	8.18	10.97
	150	7.09	8.32	11.32
Q^τ	∞	9.04	10.28	12.90
	600	9.02	10.29	13.05
	300	9.03	10.26	13.13
	150	9.15	10.42	13.32
S^μ	∞	0.348	0.461	0.745
	600	0.341	0.451	0.711
	300	0.341	0.445	0.699
	150	0.342	0.442	0.658
S^τ	∞	0.120	0.148	0.220
	600	0.118	0.145	0.209
	300	0.117	0.143	0.203
	150	0.118	0.141	0.196

Table 2. Asymptotic critical values for s_α at the ξ significance level

$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$
-0.168	0.069	0.563

Table 3. Asymptotic scaling constants for intersection of rejections testing strategies at the ξ significance level

	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$
	$IR(Q^\mu, Q^\tau)$		
m_ξ	0.801	0.793	0.782
	$IR(Q^\mu, S^\mu)$		
m_ξ^μ	0.845	0.851	0.876
	$IR(Q^\tau, S^\tau)$		
m_ξ^τ	0.897	0.894	0.900
	$IR(Q^\mu, Q^\tau, S^\mu, S^\tau)$		
m_ξ^*	0.571	0.551	0.521
	$IR(S^\mu, S^\tau)$		
m_ξ^{**}	0.576	0.554	0.522
τ_ξ	0.852	0.840	0.844

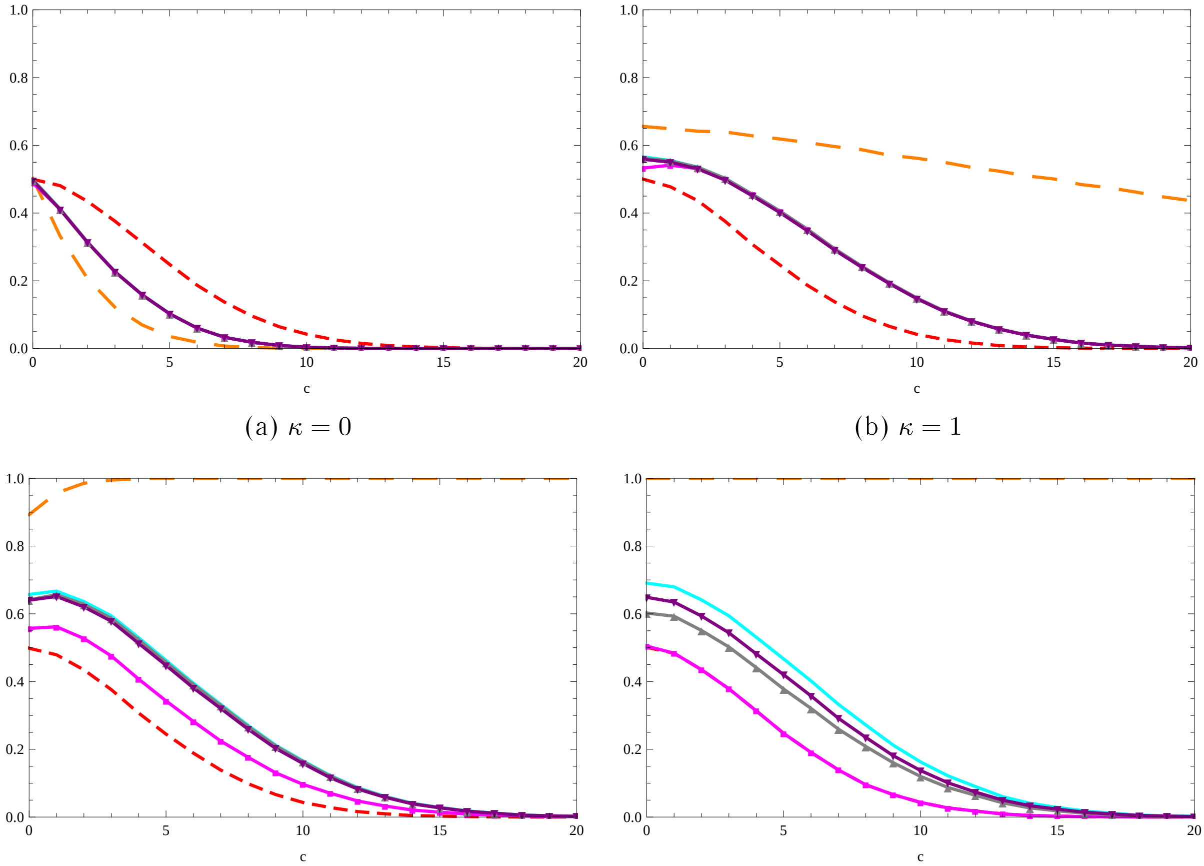


(a) $\kappa = 0$ (b) $\kappa = 0.5$

(c) $\kappa = 1$ (d) $\kappa = 2$

Figure 1. Asymptotic size and power for different values of κ, $\beta_T = \kappa\omega_\varepsilon T^{-1/2}$

Q^μ : — — , Q^τ : - - - , S^μ : — · — , S^τ : · · ·

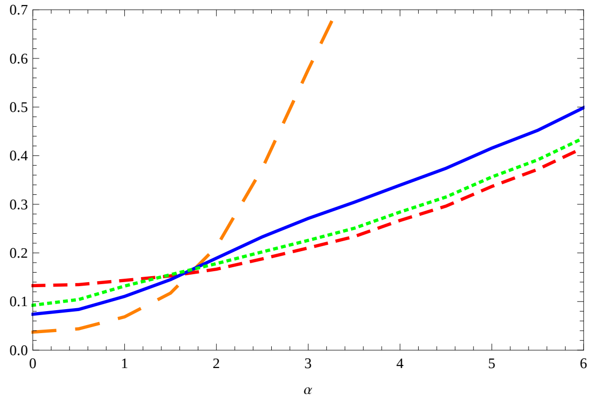


(a) $\kappa = 0$ (b) $\kappa = 1$

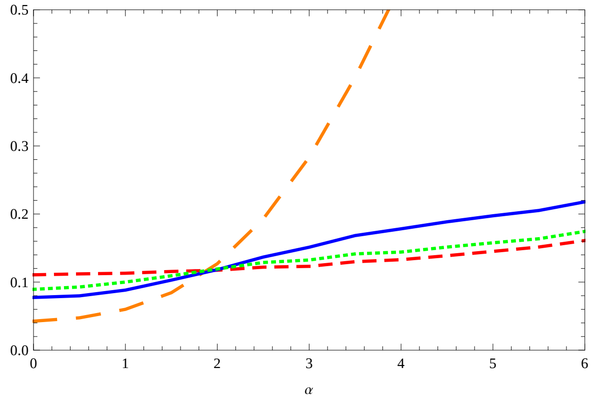
(c) $\kappa = 2$ (d) $\kappa = 4$

Figure 2. Asymptotic size and power for different values of κ, $\beta_T = \kappa\omega_\varepsilon T^{-1/2}$

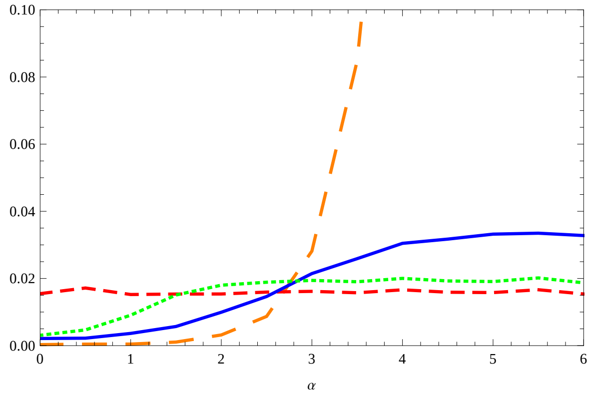
Q^{μ} : — —, Q^{τ} : - - -, IR : ——, $IR(t_\lambda)$: —■—, $IR(t_\lambda^{m2})$: —▲—, $IR(Dan\text{-}J)$: —▼—



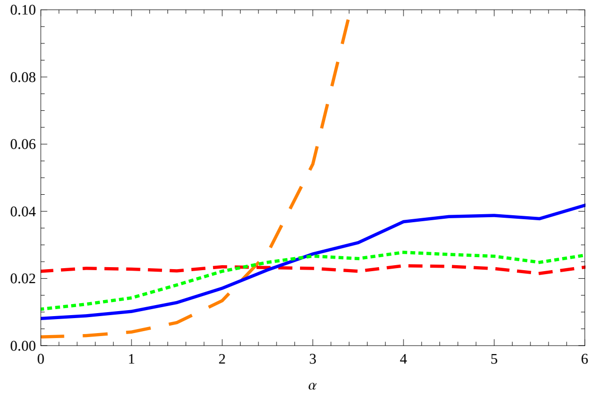
(a) $c = 5$, mean case



(b) $c = 10$, trend case



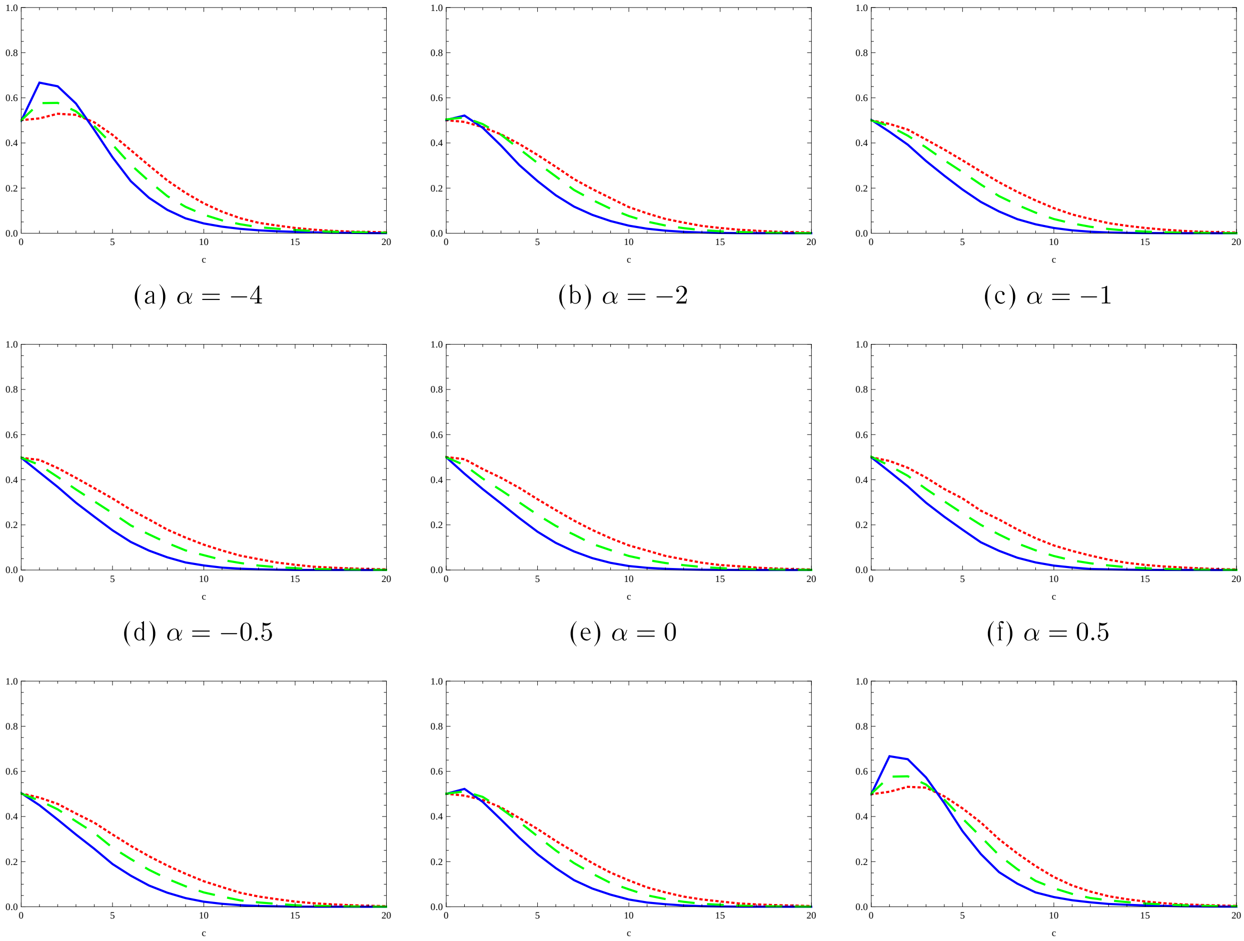
(c) $c = 10$, mean case



(d) $c = 15$, trend case

Figure 3. Asymptotic size for different initial conditions α

Q^i : — — , S^i : - - - , IR : —— , $IR(s^\alpha)$: · · ·



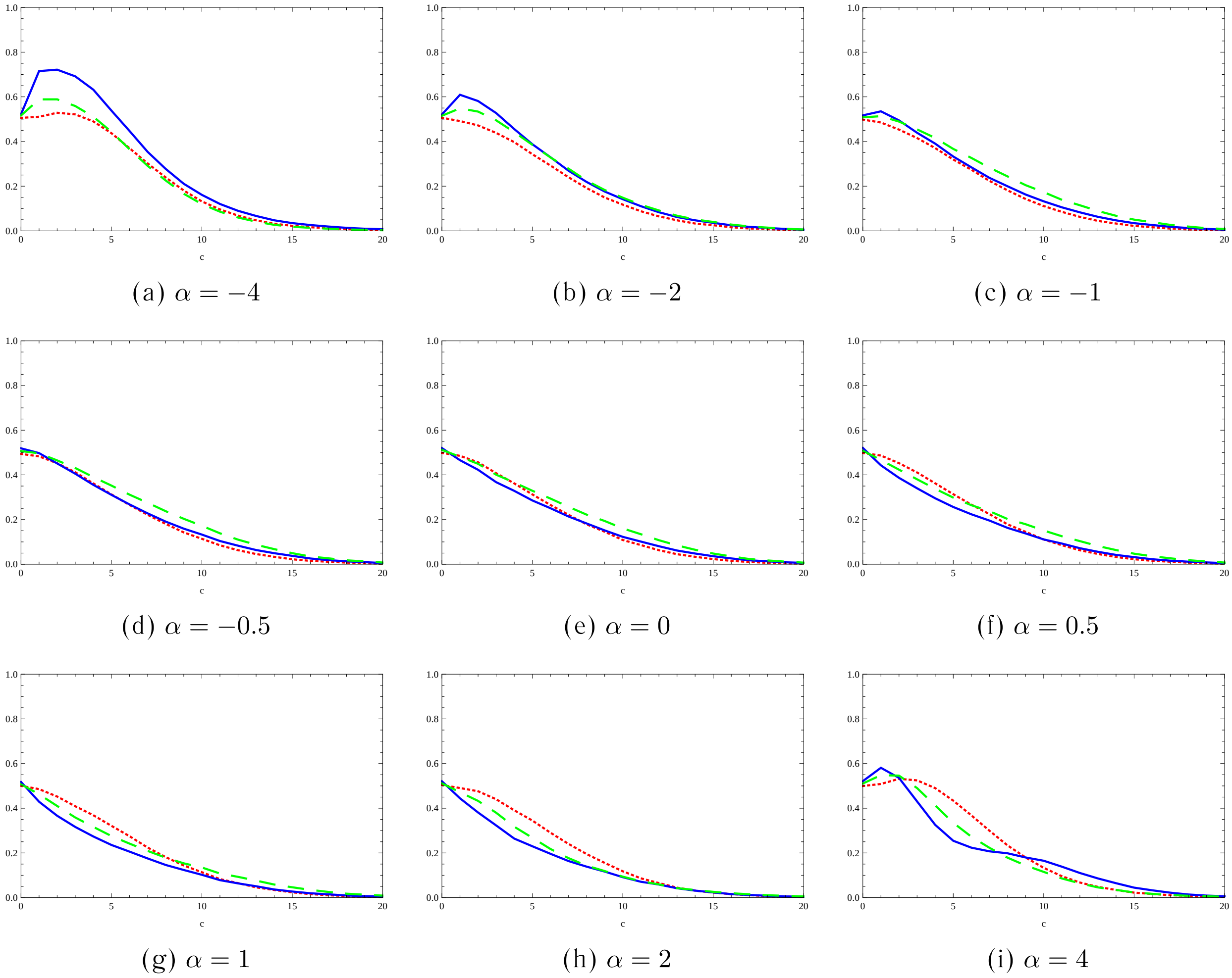
(a) $\alpha = -4$ (b) $\alpha = -2$ (c) $\alpha = -1$

(d) $\alpha = -0.5$ (e) $\alpha = 0$ (f) $\alpha = 0.5$

(g) $\alpha = 1$ (h) $\alpha = 2$ (i) $\alpha = 4$

Figure 4. Asymptotic size and power for stationarity tests, $\kappa = 0$

S^{τ} : $\cdots$, IR_4 : —— , IR_4^* : – –



(a) $\alpha = -4$ (b) $\alpha = -2$ (c) $\alpha = -1$

(d) $\alpha = -0.5$ (e) $\alpha = 0$ (f) $\alpha = 0.5$

(g) $\alpha = 1$ (h) $\alpha = 2$ (i) $\alpha = 4$

Figure 5. Asymptotic size and power for stationarity tests, $\kappa = 0.5$

S^{τ} : ⋯ , IR_4 : —— , IR_4^* : – –

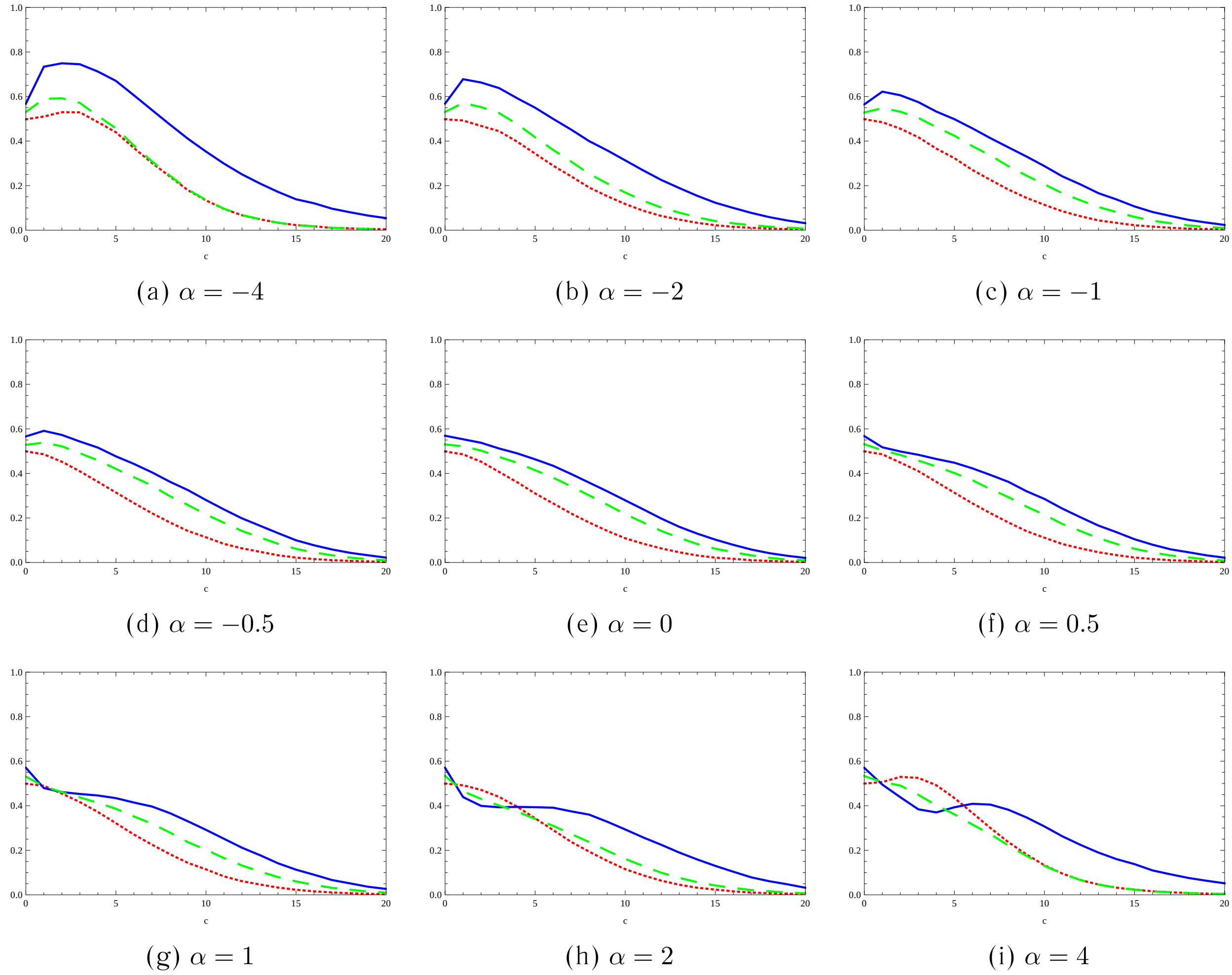


Figure 6. Asymptotic size and power for stationarity tests, $\kappa = 1$

S^{τ} : ⋯ , IR_4 : —— , IR_4^* : – –