Queen's Economics Department Working Paper No. 1431

Semiparametric Tests for the Order of Integration in the Possible Presence of Level Breaks

Fabrizio Iacone
Universita degli Studi di Milano

Morten Ørregaard Nielsen
Queen's University and CREATES

A.M. Robert Taylor
University of Essex

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

6-2020

Semiparametric Tests for the Order of Integration in the Possible Presence of Level Breaks*

Fabrizio Iacone
Università degli Studi di Milano
and University of York

Morten Ørregaard Nielsen
Queen's University
and CREATES

A.M. Robert Taylor†
University of Essex

June 15, 2020

Abstract

Lobato and Robinson (1998) develop semiparametric tests for the null hypothesis that a series is weakly autocorrelated, or $I(0)$, about a constant level, against fractionally integrated alternatives. These tests have the advantage that the user is not required to specify a parametric model for any weak autocorrelation present in the series. We extend this approach in two distinct ways. First, we show that it can be generalised to allow for testing of the null hypothesis that a series is $I(\delta)$ for any δ lying in the usual stationary and invertible region of the parameter space. Second, it is well known in the literature that long memory and level breaks can be mistaken for one another, with unmodelled level breaks rendering fractional integration tests highly unreliable. We therefore extend the Lobato and Robinson (1998) approach to allow for the possibility of changes in level at unknown points in the series. We show that the resulting statistics have standard limiting null distributions, and that the tests based on these statistics attain the same asymptotic local power functions as infeasible tests based on the unobserved errors, and hence there is no loss in asymptotic local power from allowing for level breaks, even where none is present. We report results from a Monte Carlo study into the finite-sample behaviour of our proposed tests, as well as several empirical examples.

Keywords: Fractional integration; level breaks; Lagrange multiplier testing principle; spurious long memory; local Whittle likelihood; conditional heteroskedasticity.

JEL Classification: C22

---

*Nielsen thanks the Canada Research Chairs program, the Social Sciences and Humanities Research Council of Canada, and the Center for Research in Econometric Analysis of Time Series (CREATES, funded by the Danish National Research Foundation, DNRF78) for financial support.

†Correspondence to: Robert Taylor, Essex Business School, University of Essex, Colchester, CO4 3SQ, U.K. E-mail: robert.taylor@essex.ac.uk

1 Introduction

It is well known that if not accounted for, level shifts in a weakly autocorrelated (or short memory) process, denoted $I(0)$, can induce features in the autocorrelation function and the periodogram of a time series that can be mistaken as evidence of long memory; see, *inter alia*, Diebold and Inoue (2001), Gourieroux and Jasiak (2001), Granger and Hyung (2004), Mikosch and Stărică (2004), Qu (2011), and Iacone, Leybourne, and Taylor (2019). To avoid the possibility of spurious inference being made about the memory properties of a time series, it is therefore important to develop tests on the fractional integration (memory) parameter of a time series which are robust to level shifts. As a consequence, Iacone *et al.* (2019) generalise the parametric score and Lagrange Multiplier [LM] time domain based fractional integration tests of Tanaka (1999) and Nielsen (2004) to allow for the possibility of a break in the deterministic trend function at an unknown point in the sample. These tests are equivalent to analogous extensions of the frequency domain tests of Robinson (1994) to allow for breaks in the deterministic trend function. Iacone *et al.* (2019) show that this approach delivers an LM test which, regardless of whether a break occurs or not, is a locally most powerful test and has a χ^2_1 limiting null distribution.

However, a significant practical disadvantage of the tests of Iacone *et al.* (2019) is that, like the tests of Robinson (1994), Tanaka (1999) and Nielsen (2004) from which they are derived, they are based on fitting a full parametric model to the data. Crucially, the short run component of this model must be correctly specified under the null hypothesis for the resulting test to be correctly asymptotically sized. This requirement is clearly problematic in practice, and is likely to be further complicated in the case where level breaks are present as this would likely interfere with any preliminary model selection stage used to specify the form used for the short memory component. It therefore seems worth developing long memory tests analogous to those of Iacone *et al.* (2019) but which do not require the user to specify a parametric model for the short memory component of the series.

Our contribution in this paper is therefore to develop semiparametric analogues of the parametric LM and score tests of Iacone *et al.* (2019). We will base our approach on an extension of the semiparametric frequency domain based fractional integration tests of Lobato and Robinson (1998). This approach is based on the use of a low frequency approximation provided by the local Whittle [LW] likelihood, which obviates the need to explicitly model any short range dependence present in the data. To account for the possibility of level breaks, the Lobato and Robinson (1998)-type statistics we propose are constructed from data which have been de-trended allowing for the possibility of level breaks, the locations of which are estimated by a standard residual sum of squares estimator applied to the levels data. The tests proposed in Lobato and Robinson (1998), again based on the LM testing principle, are specifically designed for testing the null hypothesis that a time series is $I(0)$. We show that, as conjectured in Lobato and Robinson

(1998, p. 478), their approach can be generalised to provide a valid test for the null hypothesis that the series is integrated of order δ, for any δ lying in the stationary and invertible region of the parameter space ($-0.5 < \delta < 0.5$). It is also possible to test orders of integration outside the stationary and invertible region using data transformations. For example, the null hypothesis of an autoregressive unit root can be obtained by testing for the null hypothesis of short memory in the first differences of the series; as such this is then a test in the levels data for a unit root allowing for the possibility of trend breaks. Because the tests are based on the LM testing principle, no preliminary estimation of the memory parameter is required.

Our focus on the Lobato and Robinson (1998) testing approach is due, at least in part, to results in Shao and Wu (2007a) who show that the standard Lobato and Robinson (1998) tests are, for a suitable choice of the bandwidth parameter m used in the local Whittle loss function, considerably more powerful than other semiparametric tests for testing the null of $I(0)$ against the alternative of fractional integration that are available in the literature. In particular, they show that tests based on the rescaled range and rescaled variance statistics and tests based on the well-known KPSS statistic of Kwiatkowski, Phillips, Schmidt and Shin (1992) have power against local alternatives of order $(\ln(T))^{-1}$, where T denotes the sample size. On the other hand, the Lobato and Robinson (1998) tests have power against local alternatives of order $m^{-1/2}$, where the bandwidth parameter m is typically of the type $m = T^{\alpha}$ for some $0 < \alpha < 1$. Moreover, these other approaches have only been developed to test the null hypothesis of $I(0)$ against the alternative of fractional integration, whereas we wish to maintain the flexibility to test a more general $I(\delta)$ null hypothesis. Busetti and Harvey (2001, 2003) develop extensions of the KPSS test that allow for a single level break at an unknown point in the sample, although their approach is based on the assumption that a level break is known to occur.

We establish that, regardless of whether level breaks occur or not, the large sample properties of the tests we propose are identical to those which obtain for the standard Lobato and Robinson (1998) tests for $\delta = 0$ in the case where no level break occurs. In particular, our proposed LM-type test has a χ^2_1 limiting null distribution and the corresponding score-based test a $N(0, 1)$ limiting null distribution, regardless of the value of δ being tested under the null hypothesis, and each attains the same asymptotic local power function as the corresponding infeasible test based on the unobserved errors. Moreover, these asymptotic local power functions do not alter between the break and no break cases and so there is no loss in asymptotic local power from allowing for level breaks, even where none is present. Although based on different and hence not directly comparable models, these large sample properties contrast with those of most popular *unit root* tests, such as that of Dickey and Fuller (1979), and *stationarity* tests, such as that of KPSS. In particular, the limiting null distributions of unit root and stationarity test statistics tend to be non-standard and depend on the functional form of the fitted deterministic, differing between the no break and break cases, and dependent on the locations of the breaks. Moreover, where breaks are fitted but not actually present in the data, these tests show a considerable decline in asymptotic local power relative to the case where no break is fitted.

The remainder of the paper is organised as follows. Section 2 sets out the fractionally integrated level break model within which we work. To aid exposition this model allows for up to one level break. Extensions to allow for multiple level breaks are discussed in Section 5. Section 3 describes our tests in the infeasible case where the errors are observable. Our proposed semiparametric statistics for the case of unknown level breaks are described in Section 4, where we also establish their large sample properties. Section 6 summarises the results from a Monte Carlo simulation study into the finite sample size and power properties of our proposed tests and compares with the nonparameteric KPSS-type tests of Busetti and Harvey (2001, 2003). Illustrative empirical examples of the methods developed in this paper to bitcoin returns data, VIX market volatility, U.S. CPI inflation, and U.S. real GDP growth are considered in Section 7. Section 8 concludes. Proofs of our main results are provided in a mathematical appendix. A supplementary appendix contains full details of the Monte Carlo design and results.

2 The Fractionally Integrated Level Break Model

Consider the scalar time series process, y_t, satisfying the data generating process (DGP),

$$y_t = \beta_1 + \beta_2 DU_t(\tau^*) + u_t, \quad t = 1, ..., T. \tag{1}$$

In (1), the level break term, $DU_t(\tau^*)$, is defined for a generic τ as $DU_t(\tau) := \mathbb{I}(t \geq \lfloor \tau T \rfloor)$, $\mathbb{I}(.)$ denoting the usual indicator function, $\lfloor . \rfloor$ denoting the integer part of its argument, and where $A := B$ and $B =: A$ is used to denote that A is defined by B. Where a level break occurs, i.e. where $\beta_2 \neq 0$, we assume that the true level break fraction is such that $\tau^* \in [\tau_L, \tau_U] =: \Lambda$, where $\Lambda \subset (0, 1)$ is compact and the quantities τ_L and τ_U are trimming parameters below and above which, respectively, a level break is deemed not to occur. In order to greatly simplify exposition and to keep notation manageable, the model considered in (1) allows for the presence of up to one level break in the data. In Section 5 we will subsequently discuss how the approach we outline in what follows for up to one level break straightforwardly extends to the case where one allows for up to k level breaks.

In the context of (1) the shocks, u_t, are assumed to follow a stationary and invertible process which is fractionally integrated of order δ, denoted $u_t \in I(\delta)$. For our purposes, we define fractional integration for u_t as

$$u_t := \Delta^{-\delta} \eta_t, \tag{2}$$

where η_t is a zero mean $I(0)$ process. We define $I(0)$ to be such that η_t has spectral density $f(\lambda)$ with $f(\lambda) \to G$ for some $G \in (0, \infty)$ as $\lambda \to 0$; formal assumptions on η_t required for our large sample theory results will be delayed until Section 3. The assumption that u_t is stationary and invertible entails that the long memory parameter, δ, is such that $\delta \in (-0.5, 0.5)$. A process satisfying the conditions just stated for u_t is often referred to in the literature as a type I fractionally integrated process.

Our interest focuses on testing the null hypothesis that u_t, and hence y_t, is $I(\delta_0)$ for some $\delta_0 \in (-0.5, 0.5)$; that is, $H_0 : \delta = \delta_0$ in (1). Based on the familiar LM testing principle we will develop tests against two-sided alternatives of the form $H_1 : \delta \neq \delta_0$ (y_t is not $I(\delta_0)$) together with corresponding score-type tests against one-sided alternatives of the form $H_1 : \delta > \delta_0$ (y_t is more persistent than an $I(\delta_0)$ series) or $H_1 : \delta < \delta_0$ (y_t is less persistent than an $I(\delta_0)$ series).

Next, in Section 3, we discuss the tests proposed in Lobato and Robinson (1998) which were developed for testing the specific null hypothesis that y_t is short memory. These tests apply to the case where either u_t in (1) is observable or where it is *known* that $\beta_2 = 0$. We show that this approach can be readily extended to develop tests for the null hypothesis that y_t is $I(\delta_0)$ for some $\delta_0 \in (-0.5, 0.5)$. Then, in Section 4, we show how these tests can be generalised to allow for the possibility that $\beta_2 \neq 0$ in (1), such that a level shift could potentially occur in the data. The testing approach we outline in Section 4 does not assume knowledge of whether a level shift genuinely occurs; that is, we do not assume knowledge of whether $\beta_2 = 0$ or $\beta_2 \neq 0$.

3 Tests of $H_0 : \delta = \delta_0$ when it is known that $\beta_2 = 0$

Suppose for the purposes of this section that it is *known* to be the case that $\beta_2 = 0$ in (1). Under this restriction we can also set $\beta_1 = 0$ with no loss of generality because, as discussed in Lobato and Robinson (1998, p. 477), the statistics we will discuss in this paper are invariant to β_1 in the case where $\beta_2 = 0$. The restriction that $\beta_2 = 0$ is therefore equivalent to the case where β_1, β_2 and τ^* are all known, such that u_t in (1) is observable. We may therefore proceed as if u_t were observable. We will discuss the application of the tests to u_t, although in the context of this section they could equally be applied to y_t because no mean-correction is required (provided the mean is constant) due to invariance to β_1.

For observable u_t, semiparametric inference on δ based on the approximation of the Whittle likelihood at low frequencies was proposed by Künsch (1987) and analysed further in Robinson (1995b). This approach is semiparametric as it does not require the specification of a parametric model for $f(\lambda)$ and, within the class of semiparametric methods, it has the advantage of being based on a (local) likelihood, and it is therefore considerably more efficient than other semiparametric estimates such as the log-periodogram regression of Geweke and Porter-Hudak (1983) and Robinson (1995a).

For a generic series a_t, let $w_a(\lambda) := \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} a_t e^{i\lambda t}$ denote the Fourier transform of a_t, and let $I_a(\lambda) := |w_a(\lambda)|^2$ denote the periodogram. Then, as discussed in Robinson (1995b), for the observable series u_t, the local Whittle estimate of δ is obtained by minimising the loss function $R(d)$ with respect to d, where

$$R(d) := \ln\left(\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d}I_u(\lambda_j)\right) - 2d\frac{1}{m}\sum_{j=1}^{m}\ln(\lambda_j) \tag{3}$$

and m denotes the bandwidth, satisfying the rate condition that $1/m + m/T \to \infty$ as $T \to \infty$. Recall that $\lambda_j := \frac{2\pi j}{T}$ for integer j are the Fourier frequencies. Applying the LM principle to the objective function in (3) yields the LM-type statistic to test $H_0 : \delta = \delta_0$,

$$LM_m^*(\delta_0) := m\left(\frac{\partial^2 R(d)}{\partial^2 d}\right)^{-1}\left(\frac{\partial R(d)}{\partial d}\right)^2\Bigg|_{d=\delta_0}.$$

Defining the score-type statistic

$$t_m^*(\delta_0) := -\left(\frac{m^{-1/2}\sum_{j=1}^{m}\nu_j\lambda_j^{2\delta_0}I_u(\lambda_j)}{\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2\delta_0}I_u(\lambda_j)}\right), \text{ where } \nu_j := \ln j - \frac{1}{m}\sum_{j=1}^{m}\ln j, \tag{4}$$

the $LM_m^*(\delta_0)$ statistic in (5) can be equivalently re-written in terms of the Fourier frequencies and the periodogram ordinates at those frequencies as

$$LM_m^*(\delta_0) = t_m^*(\delta_0)^2 = \left(\frac{m^{-1/2}\sum_{j=1}^{m}\nu_j\lambda_j^{2\delta_0}I_u(\lambda_j)}{\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2\delta_0}I_u(\lambda_j)}\right)^2. \tag{5}$$

The null hypothesis H_0 that u_t is $I(\delta_0)$ can then be rejected for large values of $LM_m^*(\delta_0)$, while a large positive (negative) value of $t_m^*(\delta_0)$ would allow rejection against the one-sided alternative $H_1 : \delta > \delta_0$ ($H_1 : \delta < \delta_0$). It will turn out that standard critical values can be employed in the context of these decision rules.

Lobato and Robinson (1998) analyse the special case of the $t_m^*(0)$ and $LM_m^*(0)$ statistics in (4) and (5), respectively, which obtain setting $\delta_0 = 0$, such that one is testing the null hypothesis of short memory, $H_0 : \delta = 0$. For the purpose of later sections, we need to also define the Lobato and Robinson (1998) score- and LM-type test statistics for the hypothesis $H_0 : \delta = \delta_0$ applied to the observed data, $\{y_t\}$, and which do not account for the possibility of level breaks; we will denote these as $t_m(\delta_0)$ and $LM_m(\delta_0)$, respectively. These differ from the infeasible statistics $t_m^*(\delta_0)$ and $LM_m^*(\delta_0)$ for the hypothesis $H_0 : \delta = \delta_0$ which are applied to the unobserved innovations, $\{u_t\}$. In the context of this section, where it is *known* that $\beta_2 = 0$, then $t_m(\delta_0)$ and $t_m^*(\delta_0)$ coincide, as do $LM_m(\delta_0)$ and $LM_m^*(\delta_0)$. Lobato and Robinson (1998) establish that, under certain regularity conditions (see Assumption 1 below), $t_m^*(0)$ and $LM_m^*(0)$ have $N(0,1)$ and χ_1^2 limiting null distributions, respectively. Shao and Wu (2007a) subsequently demonstrate that under local alternatives of the form $H_c : \delta = cm^{-1/2}$, where c is a constant (such that H_c reduces to $H_0 : \delta = 0$ when $c = 0$), $t_m^*(0) \xrightarrow{d} N(2c, 1)$ and, hence, $LM_m^*(0) \xrightarrow{d} \chi_1^2(4c^2)$, where $\chi_1^2(4c^2)$ denotes a non-central χ_1^2 distribution with non-centrality parameter $4c^2$.

Before progressing to consider the case where u_t is not observable, that is where it is not *known* for sure that $\beta_2 = 0$ in (1), we first show that the properties established for the $LM^*_m(0)$ and $t^*_m(0)$ statistics in Lobato and Robinson (1998) and Shao and Wu (2007a) carry over to the general case of the $LM^*_m(\delta_0)$ statistic in (5) and corresponding $t^*_m(\delta_0)$ statistic in (4) for testing $H_0 : \delta = \delta_0$ for any $\delta_0 \in (-0.5, 0.5)$. To do so we first introduce sufficient conditions for establishing these large sample justifications. We will discuss two sets of possible assumptions under which our large sample results obtain. The first set, given in Assumption 1, coincides with the conditions adopted by Robinson (1995b). The second set, given in Assumption 2, coincides with those employed by Shao and Wu (2007a).

Assumption 1.

(i) $\eta_t := \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ *and* ε_t *is a martingale difference sequence with* $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$, $E(\varepsilon_t^2|\mathcal{F}_{t-1}) = 1$, $E(\varepsilon_t^3|\mathcal{F}_{t-1}) = \kappa_3 < \infty$, $E(\varepsilon_t^4|\mathcal{F}_{t-1}) = \kappa_4 < \infty$, *a.s., where* $\mathcal{F}_t$ *is the* σ*-field of events generated by* ε_s, $s \leq t$.

(ii) The weights ψ_j *are such that* $\sum_{j=0}^{\infty} \psi_j^2 < \infty$.

(iii) The spectral density of η_t, $f(\lambda)$, *is twice boundedly differentiable in a neighbourhood of* $\lambda = 0$ *and satisfies, as* $\lambda \to 0^+$, *that* $f(\lambda) = G(1+O(\lambda^2))$ *and* $\frac{\partial}{\partial \lambda} \ln f(\lambda) = O(\lambda^{-1})$ *for some* $G \in (0, \infty)$.

(iv) The bandwidth, m, *is such that* $\frac{1}{m} + \frac{m^5 (\ln m)^2}{T^4} \to 0$ *as* $T \to \infty$.

Remark 3.1. The conditions on η_t detailed in Assumption 1 coincide with those given in Robinson (1995b) and are slightly stronger than those in Lobato and Robinson (1998). A full discussion of these conditions is given in Robinson (1995b, pp. 1634 and 1641) and Lobato and Robinson (1998, p. 478). Assumption 1 includes all stationary and invertible finite-order ARMA models for η_t. Assumption 1 allows for non-linearity via the martingale difference assumption on the innovations, but is otherwise linear. Notice also that Assumption 1 requires $f(\lambda)$ to be smooth only around $\lambda = 0$ and so does not rule out long memory behaviour at frequencies other than $\lambda = 0$ (although this needs to be strengthened in Assumption 3 to obtain results for our feasible tests). $\diamond$

The assumption of conditional homoskedasticity imposed by part (i) of Assumption 1 may be considered unacceptable for many data applications, in particular those involving financial data. Shao and Wu (2007a,b) show that this can be weakened to allow for a wide class of stationary, causal non-linear processes. To that end, suppose that

$$\eta_t = F(\ldots, \varepsilon_{t-1}, \varepsilon_t), \tag{6}$$

where ε_t are independent and identically distributed (IID) random variables and F is a measurable function such that η_t is well defined as a stationary, causal, ergodic process. For a random variable ξ and $p > 0$, write $\xi \in \mathcal{L}^p$ if $\|\xi\|_p := (E(|\xi|^p))^{1/p} < \infty$. Let $\{\varepsilon'_t\}$,

$t \in Z$, be an IID copy of $\{\varepsilon_t\}$, $\mathcal{F}_t = (\ldots, \varepsilon_{t-1}, \varepsilon_t)$, $\mathcal{F}_k^* := (\mathcal{F}_{-1}, \varepsilon_0)$, $\eta_k' = F(\mathcal{F}_0^*, \varepsilon_1, \ldots, \varepsilon_t)$ and $\vartheta_q(k) = \|\eta_k - \eta_k'\|_q$.

Assumption 2. *For η_t and F defined as in (6) and for some $q > 4$:*

(i) $\eta_t \in \mathcal{L}^q$ *and* $\sum_{k_1,k_2,k_3} cum(\eta_0, \eta_{k_1}, \eta_{k_2}, \eta_{k_3}) < \infty$.

(ii) $\sum_{k=1}^{\infty} k\vartheta_q(k) < \infty$.

(iii) The spectral density of η_t, $f(\lambda)$, satisfies $f(\lambda) = G(1 + O(\lambda^2))$ as $\lambda \to 0^+$ for some $G \in (0, \infty)$.

(iv) The bandwidth, m, is such that $\frac{(\ln T)^3}{m} + \frac{m}{T^{2/3}} \to 0$ as $T \to \infty$.

Remark 3.2. Assumption 2 includes a number of widely used nonlinear time series models for η_t such as bilinear models, threshold models, GARCH and ARMA-GARCH models; see Shao and Wu (2007a, p. 254) and Shao and Wu (2007b) and the references therein for further discussion of this assumption and further examples of classes of nonlinear processes which satisfy it. While Assumption 2 weakens, *inter alia*, the conditional homoskedasticity restriction of Assumption 1, this comes at the cost of a stronger assumption on the bandwidth, that is restricted to be such that $m = o(T^{2/3})$. Moreover, as discussed in Shao and Wu (2007b, Remark 3.1), Assumption 2(ii) implies continuous differentiability of $f(\lambda)$ for all frequencies, whereas, as discussed in Remark 3.1 and Robinson (1995b), Assumption 1 only imposes conditions on $f(\lambda)$ in a local-to-zero band. There is therefore a clear trade-off between the conditions imposed on η_t by Assumptions 1 and 2. ◊

In Theorem 1 we now derive the large sample properties of the $LM_m^*(\delta_0)$ and $t_m^*(\delta_0)$ statistics, obtained for the case where it is *known* that $\beta_2 = 0$ in (1). To facilitate discussion of asymptotic local power, we consider the local alternative $H_c : \delta = \delta_0 + cm^{-1/2}$.

Theorem 1. *Let y_t be generated according to (1) with $\beta_2 = 0$, and let either Assumption 1 or Assumption 2 hold on η_t. Then, for any $\delta_0 \in (-0.5, 0.5)$, under $H_c : \delta = \delta_0 + cm^{-1/2}$:*

(i) $LM_m^*(\delta_0) \xrightarrow{d} \chi_1^2(4c^2)$; *and*

(ii) $t_m^*(\delta_0) \xrightarrow{d} N(2c, 1)$.

Remark 3.3. Theorem 1 shows that the results obtained for the limiting null distributions of the $LM_m^*(0)$ and $t_m^*(0)$ statistics in Lobato and Robinson (1998) apply more generally to the $LM_m^*(\delta_0)$ and $t_m^*(\delta_0)$ statistics for testing the null hypothesis that u_t is $I(\delta_0)$ for any δ_0 in the usual stationary and invertible region. Theorem 1 also shows that tests based on the $LM_m^*(\delta_0)$ and $t_m^*(\delta_0)$ statistics posses the same local power functions as tests based on the $LM_m^*(0)$ and $t_m^*(0)$ statistics. Moreover, these results hold regardless of whether u_t is conditionally homoskedastic or conditionally heteroskedastic (satisfying Assumption 2). ◊

4 Feasible Tests of $H_0 : \delta = \delta_0$ Allowing for a Level Break

Recall that the LM- and score-based tests discussed in Section 3 are based on the assumption that $\beta_2 = 0$, such that the $LM_m(\delta_0)$ and $t_m(\delta_0)$ statistics calculated on the observed data $\{y_t\}$ will coincide with the $LM_m^*(\delta_0)$ and $t_m^*(\delta_0)$ statistics based on the shocks, $\{u_t\}$, even if $\beta_1 \neq 0$ such that $\{u_t\}$ are unobservable (because the statistics are invariant to β_1). However, where $\beta_2 \neq 0$ this is no longer the case, and we cannot proceed as if the tests were based on the unobservable shocks, $\{u_t\}$. Moreover, where $\beta_2 \neq 0$ the $LM_m(\delta_0)$ and $t_m(\delta_0)$ statistics constructed from the observed data, $\{y_t\}$, are non-similar tests and will diverge. For example, if $\delta_0 = 0$ it can be shown that the (exact) rates of divergence are $LM_m(0) = O_p(m \ln(m)^2)$ and $t_m(0) = O_p(\sqrt{m} \ln(m))$ under H_0, so that both statistics will diverge with the sample size, even under the null hypothesis. As a consequence, therefore, the Lobato and Robinson (1998) tests will spuriously reject the null with probability tending to one as the sample size diverges. That is, tests based on $LM_m(\delta_0)$ or $t_m(\delta_0)$ are uninformative if it is unknown whether $\beta_2 = 0$ or not. In this section we will therefore discuss how feasible versions of the tests discussed in Section 3 can be derived for the case where it is not known for certain whether $\beta_2 = 0$ or not.

In the context of (1), the disturbances u_t are not observable and so they must be estimated. For a generic (putative) break location, τ, we can do so by using standard ordinary least squares (OLS) estimators of the parameters β_1 and β_2 in (1). To that end, let $\beta := (\beta_1, \beta_2)'$, and let $y := (y_1, \dots, y_T)'$, $x_t(\tau) := (1, DU_t(\tau))'$, and $x(\tau) := (x_1(\tau), \dots, x_T(\tau))'$. Then the OLS estimate of β is given by $\widehat{\beta}(\tau) = (\widehat{\beta}_1(\tau), \widehat{\beta}_2(\tau))' := (x(\tau)'x(\tau))^{-1}x(\tau)'y$. For a given value of τ we then have the corresponding estimated residuals $\widehat{u}_t(\tau) := y_t - \widehat{\beta}(\tau)'x(\tau)$, with associated periodogram $I_{\widehat{u}(\tau)}(\lambda_j)$.

Based on $I_{\widehat{u}(\tau)}(\lambda_j)$, we can then define analogues of the $LM_m^*(\delta_0)$ statistic of (5) and the corresponding score-based statistic $t_m^*(\delta_0)$ in (4), for testing $H_0 : \delta = \delta_0$ as follows

$$t_m(\delta_0; \tau) := -\left(\frac{m^{-1/2} \sum_{j=1}^{m} \nu_j \lambda_j^{2\delta_0} I_{\widehat{u}(\tau)}(\lambda_j)}{\frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2\delta_0} I_{\widehat{u}(\tau)}(\lambda_j)} \right) \tag{7}$$

$$LM_m(\delta_0; \tau) := (t_m(\delta_0; \tau))^2. \tag{8}$$

If the true break fraction, τ^*, were known then one would simply evaluate $LM_m(\delta_0; \tau)$ and $t_m(\delta_0; \tau)$ at $\tau = \tau^*$. Our focus, however, is on the case where τ^* is unknown and so will need to be estimated from the data. An obvious candidate is the minimum residual sum of squares [RSS] estimator considered in Iacone *et al.* (2019), viz,

$$\widehat{\tau} := \arg \min_{\tau \in [\tau_L, \tau_U]} \sum_{t=1}^{T} \widehat{u}_t(\tau)^2, \tag{9}$$

where it is recalled that τ_L and τ_U are trimming parameters such that $[\tau_L, \tau_U] \subset (0, 1)$. Tests for $H_0 : \delta = \delta_0$ could then be based on $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$.

For these tests to be operational, we will need to establish the large sample behaviour of the $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ statistics under the null hypothesis, $H_0 : \delta = \delta_0$, and show that unique asymptotic critical values (in the sense that they do not depend on any nuisance parameters) for the tests can be obtained from these distributions. In fact, we will be able to show in what follows that these statistics have the same limiting null distributions as were obtained for their infeasible counterparts $LM_m^*(\delta_0)$ and $t_m^*(\delta_0)$ in Theorem 1. In order to do so, however, we must impose some additional regularity conditions on η_t. In particular, Assumptions 1 and 2 must be strengthened to Assumptions 3 and 4, respectively, as follows:

Assumption 3. *Let Assumption 1 hold. Assume further that:*

(i) $\sup_t E(|\varepsilon_t|^q) < \infty$ *for some* $q > 1/(1 + 2\delta)$.

(ii) The weights ψ_j *are such that* $\sum_{j=0}^{\infty} j|\psi_j| < \infty$.

(iv) For some $\epsilon > 0$*, the bandwidth,* m*, is such that* $\frac{T^\epsilon}{m} \to 0$ *as* $T \to \infty$.

Assumption 4. *Let Assumption 2 hold, and define the projection operator* $\wp_k \xi := E(\xi|\mathcal{F}_k) - E(\xi|\mathcal{F}_{k-1})$. *Then we assume further that:*

(i) $\eta_t \in \mathcal{L}^q$ *and* $\sum_{k=0}^{\infty} \|\wp_0 \eta_k\|_q < \infty$ *for some* $q > 1/(1 + 2\delta)$.

(iv) For some $\epsilon > 0$*, the bandwidth,* m*, is such that* $\frac{T^\epsilon}{m} \to 0$ *as* $T \to \infty$.

Remark 4.1. Both Assumptions 3 and 4 impose the additional moment condition that $q > 1/(1 + 2\delta)$ moments exist. This condition is needed so that we can appeal to the functional central limit theorem [FCLT] for fractional processes for which the moment condition is necessary; see Theorem 2 of Johansen and Nielsen (2012). The fractional FCLT also requires that $q > 2$, but this is implied in Assumptions 1 or 2 so is not stated explicitly here. The condition placed on the weights ψ_j in Assumption 3(ii) is quite standard for the (fractional) FCLT and is met by all stationary and invertible finite-order ARMA models. This condition also implies continuity of the spectral density of η_t and hence rules out long memory at other frequencies, see Remarks 3.1 and 3.2. The condition that $0 < \sum_{j=0}^{\infty} \psi_j < \infty$ (and a similar condition for the non-linear process) is again omitted because it is implied by the assumption $0 < f(0) < \infty$. The additional condition required to hold on the bandwidth in part (iv) of Assumptions 3 and 4 is not restrictive in practice because much larger bandwidths will typically be used. $\diamondsuit$

We are now in a position to state our main result in Theorem 2 which details the large sample behaviour of the feasible statistics $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ under local alternatives of the form $H_c : \delta = \delta_0 + cm^{-1/2}$. We will do this by comparing them to the infeasible $LM_m^*(\delta_0)$ and $t_m^*(\delta_0)$ statistics, respectively. Inherent in doing so will be to analyse the distance between $\widehat{u}_t$ and $\widehat{u}_t(\widehat{\tau})$, the latter given by $\widehat{u}_t(\tau)$ evaluated at $\tau = \widehat{\tau}$, and establish how this affects the distance between the feasible and infeasible statistics. The behaviour of both $LM_m(\delta_0, \widehat{\tau})$ and $t_m(\delta_0, \widehat{\tau})$ clearly depend on the large sample properties of the estimates $\widehat{\tau}$ of (9) and $\widehat{\beta}(\widehat{\tau})$, the latter given by $\widehat{\beta}(\tau)$ evaluated at $\tau = \widehat{\tau}$. For the properties of $\widehat{\tau}$ we apply a result of Lavielle and Moulines (2000), and we combine this with a fractional FCLT for u_t to obtain results for $\widehat{\beta}(\tau)$ following a similar analysis to that in Lemma 1 of Iacone *et al.* (2019). In Theorem 2 we then establish that these properties are sufficient to show that the differences, $LM_m(\delta_0, \widehat{\tau}) - LM_m^*(\delta_0)$ and $t_m(\delta_0, \widehat{\tau}) - t_m^*(\delta_0)$, are both asymptotically negligible, regardless of whether or not a trend break occurs (i.e., whether $\beta_2 \neq 0$ or $\beta_2 = 0$).

Theorem 2. *Let y_t be generated according to (1), and let either Assumption 3 or Assumption 4 hold on η_t. Then, for any $\delta_0 \in (-0.5, 0.5)$, under $H_c : \delta = \delta_0 + cm^{-1/2}$:*

(i) If $\beta_2 = 0$, then $t_m(\delta_0; \tau) - t_m^*(\delta_0) = o_p(1)$ and $LM_m(\delta_0; \tau) - LM_m^*(\delta_0) = o_p(1)$, in each case uniformly in τ.

(ii) If $\beta_2 \neq 0$, then $t_m(\delta_0; \widehat{\tau}) - t_m^*(\delta_0) = o_p(1)$ and $LM_m(\delta_0; \widehat{\tau}) - LM_m^*(\delta_0) = o_p(1)$.

Some remarks are in order.

Remark 4.2. An immediate consequence of Theorem 2 is that both $LM_m(\delta_0; \widehat{\tau}) - LM_m^*(\delta_0)$ and $t_m(\delta_0; \widehat{\tau}) - t_m^*(\delta_0)$ are of $o_p(1)$ irrespective of whether $\beta_2 \neq 0$ or $\beta_2 = 0$. Consequently, regardless of the value of β_2, $LM_m(\delta_0; \widehat{\tau}) \xrightarrow{d} \chi_1^2(4c^2)$ and $t_m(\delta_0; \widehat{\tau}) \xrightarrow{d} N(2c, 1)$ under H_c, and tests based on these statistics thereby attain exactly the same asymptotic local power functions as obtained by the infeasible tests based on $LM_m^*(\delta_0)$ and $t_m^*(\delta_0)$, respectively. Moreover, since $LM_m(\delta_0; \widehat{\tau}) \xrightarrow{d} \chi_1^2$ and $t_m(\delta_0; \widehat{\tau}) \xrightarrow{d} N(0, 1)$ under H_0, standard critical values can be used. ♢

Remark 4.3. The result given in part (i) of Theorem 2 demonstrates that when no level break occurs, the difference between the statistics based on $\widehat{u}_t$ and $\widehat{u}_t(\tau)$ is asymptotically negligible, and that this holds uniformly in τ and, hence, holds for $\widehat{\tau}$. Part (ii) of Theorem 2 shows that when $\beta_2 \neq 0$, such that a level break occurs, the differences between the semiparametric statistics based on $\widehat{u}_t$ and $\widehat{u}_t(\widehat{\tau})$ are asymptotically negligible. This arises because $\widehat{\tau} \xrightarrow{p} \tau^*$ at a sufficiently fast rate; cf. part (i) of Lemma 1 in Iacone *et al.* (2019). ♢

Remark 4.4. A formal proof of Theorem 2 is provided in the appendix. However, it seems useful to summarise the essential parts of the proof here. Beginning with the case in which in fact no level break occurs ($\beta_2 = 0$), we first establish uniformly in τ results for $\widehat{\beta}(\tau)$. It is at this stage that the fractional FCLT is used. We can then derive properties of the estimated residuals $\widehat{u}_t(\tau)$ and analyse the distance between the Fourier transforms (and hence the periodograms) of $\widehat{u}_t(\tau)$ and of u_t. In the second part of the proof we consider the case in which a break does occur ($\beta_2 \neq 0$). In this case, we first establish the properties of the estimate of the break, $\widehat{\tau}$, using results from Lavielle and Moulines (2000). These properties allow us to bound the distance between $\widehat{\beta}(\widehat{\tau})$ and $\widehat{\beta}(\tau^*)$, and use this to study the distance between the Fourier transforms (and the periodograms) of $\widehat{u}_t(\widehat{\tau})$ and of $\widehat{u}_t(\tau^*)$. $\diamondsuit$

5 Multiple Level Breaks

In this section we briefly discuss how the single level break model in (1) can be extended to allow for the (possible) presence of multiple level breaks. As we will show, the resulting semiparametric LM- and score-type statistics still satisfy the same large sample properties given in Theorem 2 for the case of up to one level break.

To that end, we can generalise (1) to allow for (up to) k level breaks, viz,

$$y_t = \beta_1 + \boldsymbol{\beta}_2' \boldsymbol{DU}_t(\boldsymbol{\tau}^*) + u_t,$$

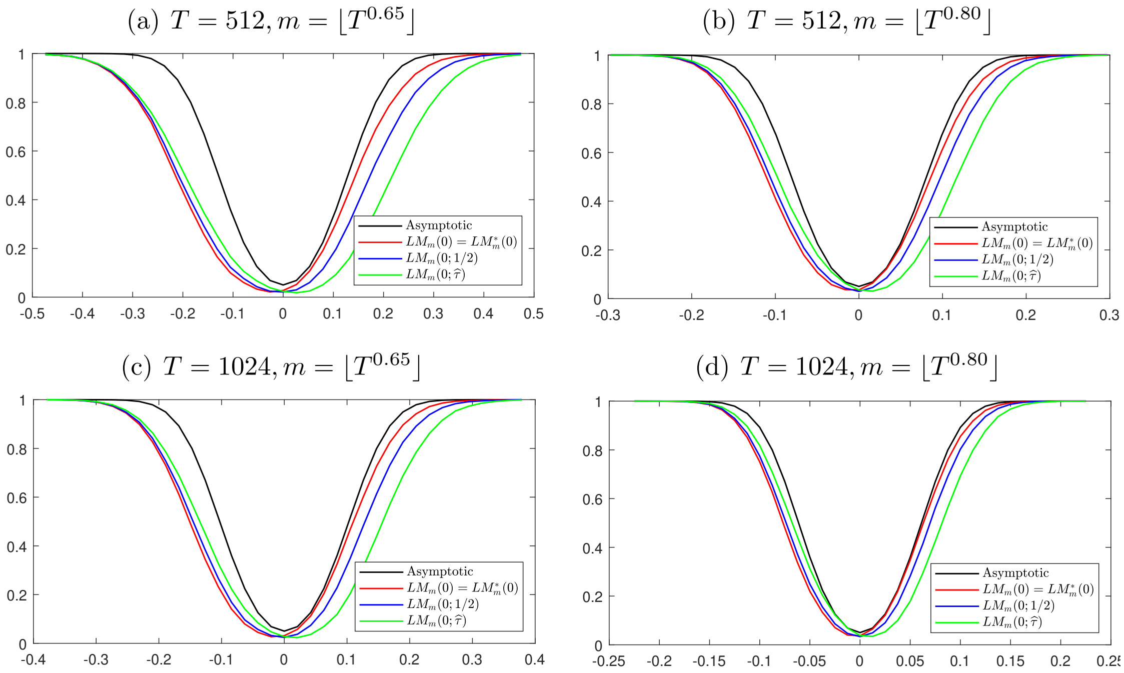
where $\boldsymbol{DU}_t(\boldsymbol{\tau}^*) := (DU_t(\tau_1^*), \ldots, DU_t(\tau_k^*))'$. Here $\boldsymbol{\tau}^* := (\tau_1^*, \ldots, \tau_k^*)'$ is the vector of (unknown) putative level break fractions and $\boldsymbol{\beta}_2 := (\beta_{2,1}, \ldots, \beta_{2,k})'$ the associated break magnitude parameters, such that a level break occurs at time $\lfloor \tau_i^* T \rfloor$ when $\beta_{2,i} \neq 0$ for $i = 1, \ldots, k$. The break fractions are assumed to be such that $\tau_i^* \in \Lambda$ for all $i = 1, \ldots, k$. A standard assumption in such a model is that $|\tau_i^* - \tau_j^*| \geq \lambda > 0$ for all $i \neq j$, such that the DGP admits (up to) k level breaks occurring at unknown points across the interval Λ, with at least $\lfloor \lambda T \rfloor$ observations between breaks (note that k and λ must satisfy the relation $k \leq 1 + \lfloor (\tau_U - \tau_L)/\lambda \rfloor$).

Provided that the k possible break locations are estimated using the obvious k-dimensional analogue of (9), a precise formula for which is given in (16) of Lavielle and Moulines (2000, p. 38), yielding the vector of estimates, $\widehat{\boldsymbol{\tau}} := (\widehat{\tau}_1, \ldots, \widehat{\tau}_k)'$ say, then it can be shown that the corresponding semiparametric statistics, $LM_m(\delta_0, \widehat{\boldsymbol{\tau}})$ and $t_m(\delta_0, \widehat{\boldsymbol{\tau}})$ say, will have precisely the same properties as $LM_m(\delta_0, \widehat{\tau})$ and $t_m(\delta_0, \widehat{\tau})$, respectively, in Theorem 2. That is, $LM_m(\delta_0, \widehat{\boldsymbol{\tau}}) \xrightarrow{d} \chi_1^2(4c^2)$ and $t_m(\delta_0, \widehat{\boldsymbol{\tau}}) \xrightarrow{d} N(2c, 1)$ under H_c, and $LM_m(\delta_0, \widehat{\boldsymbol{\tau}}) \xrightarrow{d} \chi_1^2$ and $t_m(\delta_0, \widehat{\boldsymbol{\tau}}) \xrightarrow{d} N(0, 1)$ under H_0, in both cases irrespective of whether $\beta_{2,i} = 0$ or $\beta_{2,i} \neq 0$ for any particular $i = 1, \ldots, k$. The demonstration of this result is a straightforward but tedious extension of the proof of the results stated in Theorem 2 for the $k = 1$ case. In particular, the result follows using the key results that: (a) for those of the k locations where a break does occur, i.e. for those values of $i = 1, \ldots, k$ such that $\beta_{2,i} \neq 0$, Lavielle and Moulines (2000) demonstrate that $\widehat{\tau}_i \xrightarrow{p} \tau_i^*$ and at the same rate as $\widehat{\tau} \xrightarrow{p} \tau^*$ in the $k = 1$ case, and (b) for the remaining locations where no break occurs, i.e. for those values of $i = 1, \ldots, k$ such that $\beta_{2,i} = 0$, analogous uniformity arguments to those made in the proof of Theorem 2 for the $k = 1$ case hold.

6 Monte Carlo Evidence

We begin this section by investigating how well the large sample predictions of Theorem 2 hold in finite samples. To that end, Figures 1 and 2 graph simulated finite sample power functions of the the feasible LM-type test, $LM_m(\delta_0; \widehat{\tau})$, proposed in Section 4 and the

Figure 1: Rejection frequencies, $\delta_0 = 0$ and $\beta_2 = 0$



corresponding Lobato and Robinson (1998) test, $LM_m(\delta_0)$, that does not allow for the possibility of a level break. In the context of the $LM_m(\delta_0; \widehat{\tau})$ statistic we set the trimming parameters to be $\tau_L = 0.15$ and $\tau_U = 0.85$. Also graphed are the power functions of the corresponding infeasible tests, $LM_m(\delta_0; \tau^*)$, defined just under (8), and $LM_m^*(\delta_0)$ defined in (5). The former assumes knowledge of the true break location, τ^*, but not the innovations, u_t, and the latter assumes knowledge of the innovations.

The simulated data used to construct the power curves in Figures 1 and 2 were generated according to the DGP in (1)–(2) for $T = 512$ and $T = 1024$ with $\eta_t \sim NIID(0, 1)$, and where β_1 was set equal to zero with no loss of generality. All of the reported tests are for testing $H_0 : \delta = 0$ at the nominal asymptotic 5% level. The graphs depict the simulated power functions of the tests under the local alternative $H_c : \delta = cm^{-1/2}$ for a range of values of c and with the corresponding values of δ shown on the horizontal axes. Results are reported for two bandwidth choices, namely $m = \lfloor T^{0.65} \rfloor$ and $m = \lfloor T^{0.8} \rfloor$. The results in Figure 1 relate to the case considered in part (i) of Theorem 2 with no level break, i.e. $\beta_2 = 0$, while the results in Figure 2 relate to part (ii) of Theorem 2 for the specific case of a level break with $\beta_2 = 2$ at $\tau^* = 0.5$, i.e. a break equal to two standard deviations of the innovation process occurring midway through the sample. The simulated power curves were computed using 10,000 Monte Carlo replications using the `RNDN` function of Gauss 20. As a benchmark, we also include in each graph the corresponding asymptotic local power curves obtained directly from the non-central $\chi_1^2(4c^2)$ distribution, where $c = \delta\sqrt{m}$.

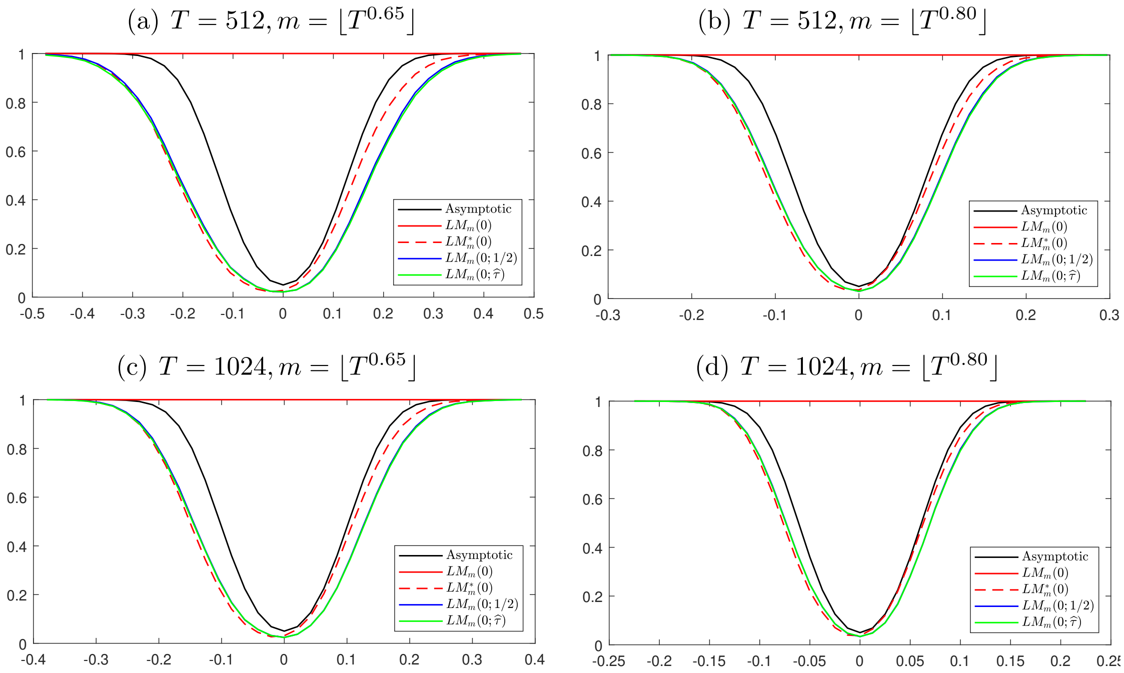
Consider first the results in Figure 1 for the no break case. Here, given knowledge that no level break was present, the best possible test to use among the three considered would be the basic Lobato and Robinson (1998) test, $LM_m(\delta_0) = LM^*_m(\delta_0)$. Against positive values of δ this test has power closest to the asymptotic local power function and is somewhat more powerful than the infeasible $LM_m(\delta_0; \tau^*)$ test, which in turn is more powerful than the feasible $LM_m(\delta_0; \widehat{\tau})$ test. These differences are, however, reduced for $T = 1024$ *vis-à-vis* $T = 512$ and for $m = \lfloor T^{0.8} \rfloor$ *vis-à-vis* $m = \lfloor T^{0.65} \rfloor$; indeed for $T = 1024$ and $m = \lfloor T^{0.8} \rfloor$ the differences between the three tests are quite small with all three lying close to the asymptotic local power curve. For negative values of δ there are only very slight differences between the three tests. Overall, the large sample predictions from part (i) of Theorem 2 appear to hold reasonably well in finite samples, particularly so for the larger bandwidth considered.

Consider next the the results in Figure 2 for the case where a level break of magnitude $\beta_2 = 2$ occurs. Here the infeasible $LM^*_m(\delta_0)$ test no longer coincides with the feasible Lobato and Robinson (1998) test, $LM_m(\delta_0)$. In this case the divergence of the $LM_m(\delta_0)$ test is clearly seen, regardless of whether the null hypothesis holds or not, with the test rejecting essentially 100% of the time even for the smaller sample size considered. The power functions of the infeasible $LM_m(\delta_0; \tau^*)$ and feasible $LM_m(\delta_0; \widehat{\tau})$ tests essentially coincide regardless of the sample size or bandwidth considered, suggesting that τ^* is very accurately estimated by $\widehat{\tau}$ in this case. As with the results for the no break case in Figure 1, for positive values of δ the power curve of the feasible $LM_m(\delta_0; \widehat{\tau})$ test lies only slightly below that of the infeasible $LM^*_m(\delta_0)$ test, which in turn lies close to the asymptotic local power curve, with the differences between the power curves reducing as T and/or m is increased. For negative values of δ the power curves of the $LM_m(\delta_0; \widehat{\tau})$ and $LM^*_m(\delta_0)$ tests are almost indistinguishable regardless of m or T. Again the large sample predictions from part (ii) of Theorem 2 would appear to hold reasonably well in finite samples.

In the remainder of this section we summarise the results from an large set of Monte Carlo experiments designed to investigate the finite sample size and power properties of the semiparametric long memory tests proposed in Section 4. Specifically, we compare the empirical size and power properties of the $LM_m(\delta_0; \widehat{\tau})$, $LM_m(\delta_0; \tau^*)$ and $LM_m(\delta_0)$ tests featured in Figures 1 and 2 along with the corresponding t-type tests, $t_m(\delta_0; \widehat{\tau})$, $t_m(\delta_0; \tau^*)$ and $t_m(\delta_0)$, respectively. Comparison is also made with the KPSS stationarity test, denoted $KPSS$, together with the generalisations thereof proposed in Busetti and Harvey (2001, 2003) which allow for a level break at either a known or unknown location, denoted $KPSS(\tau^*)$ and $KPSS(\widehat{\tau})$, respectively. The full set of results together with details of the experimental design can be found in the supplementary appendix.

We considered DGPs for $\{y_t\}$ of the form given in (1) testing $H_0 : \delta = 0$; both where $\delta = 0$ (empirical size) and where $\delta \in \{-0.15, 0.15\}$ (empirical power). The empirical size properties of tests for $H_0 : \delta = 0.3$ and $H_0 : \delta = -0.3$ were also explored. The DGP had either no level break or a level break at the sample midpoint with magnitude $\beta_2 \in$

Figure 2: Rejection frequencies, $\delta_0 = 0$ and $\beta_2 = 2$



$\{0.5, 1, 2\}$. For the empirical size results the error process η_t was allowed to follow either an IID process, an AR(1) process or an ARCH(1) process, while for empirical power IID and ARCH(1) processes were considered. Again we set the search set as $\Lambda = [0.15, 0.85]$. All of the tests were implemented for both a range of fixed bandwidths and using data-driven bandwidth rules. The principal findings of our Monte Carlo results can be summarised as follows:

(i) As with the findings in Lobato and Robinson (1998) our results demonstrate that the bandwidth m has a significant impact on the finite sample properties of the tests, with a clear trade-off seen between size and power. In particular, for a given sample size, excluding those tests which are non-similar (i.e., excluding the $LM_m(\delta_0)$ and $t_m(\delta_0)$ tests when $\beta_2 \neq 0$), we observe the following general patterns: (a) for a given pattern of weak dependence and a given bandwidth, m, the observed distortions from the nominal (asymptotic) significance level are greater the larger is m, and (b) empirical power against a given fixed alternative increases as the bandwidth, m, increases. Generally, a range of bandwidths between $m = \lfloor T^{0.5} \rfloor$ and $m = \lfloor T^{0.65} \rfloor$ provides reasonable finite sample size control across the cases considered.

(ii) Our results suggest that the automatic bandwidth, m_{LR}, of Lobato and Robinson (1998) delivers a reasonable trade-off between finite sample size and power considerations, at least when the data are conditionally homoskedastic. In the conditionally heteroskedastic ARCH(1) case, the empirical size of tests based on m_{LR} do not improve, other things equal, as the sample size is increased. This is perhaps not surprising given that the m_{LR} bandwidth rule is not consistent with the bandwidth rate imposed on m by Assumption 2, and we therefore recommend caution in using the m_{LR} bandwidth rule with data which are suspected to display conditional heteroskedasticity. For the KPSS-type tests, the automatic bandwidth rule recommended in Lobato and Robinson (1998) also appears to deliver a reasonable size-power trade-off.

(iii) Overall, our results suggest that it may be helpful in practice to consider the automatic bandwidth, m_{LR}, together with a range of bandwidths between $m = \lfloor T^{0.5} \rfloor$ and $m = \lfloor T^{0.65} \rfloor$. This is what we will do in the empirical examples in Section 7.

(iv) As expected, where a level break occurs ($\beta_2 \neq 0$), the non-similar $LM_m(\delta_0)$, $t_m(\delta_0)$ and $KPSS$ tests are highly unreliable displaying severe oversize (excepting the left-tailed $t_m(\delta_0)$ test which is correspondingly undersized), and hence spurious evidence of long memory. The observed size distortions seen with these tests are higher, other things equal, the larger is the sample size or the level break magnitude.

(v) Although asymptotically equivalent under both the null and local alternatives (cf. Theorem 2), differences are observed between the finite sample size and power properties of the pairs of tests $LM_m(\delta_0; \widehat{\tau})$ and $LM_m(\delta_0; \tau^*)$, and $t_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \tau^*)$. The $LM_m(\delta_0; \tau^*)$ and $t_m(\delta_0; \tau^*)$ tests are based on knowledge of whether a level break occurs or not (i.e. whether $\beta_2 = 0$ or $\beta_2 \neq 0$) and, where a break occurs, also knowledge of the level break location τ^*, while $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ do not assume knowledge of either. The differences between the finite sample properties of these pairs of tests are seen to diminish as either the sample size or, in the case where a level break occurs, the break magnitude increases; indeed, for the largest magnitude considered, $\beta_2 = 2$, these differences are largely eliminated even for the smaller of the two sample sizes considered. The observed differences between the empirical power properties of these pairs of tests are seen to be slightly larger, other things equal, in the case where the errors are ARCH(1) *vis-à-vis* the IID case. Moreover, the finite sample differences between the pairs of tests are smallest for the tests of $H_0 : \delta = -0.3$ and largest for the tests of $H_0 : \delta = 0.3$. Where no level break is present, the finite sample differences between the $LM_m(\delta_0; \widehat{\tau})$ test and $LM_m(\delta_0)$ (which assume no level break is present) are again relatively small, other things equal, particularly for the larger sample size considered. This is also broadly true for a comparison between the $t_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0)$ tests, although the differences are larger than for the LM-type tests. Overall, the asymptotic theory presented in Theorem 2 appears to provide a reasonable prediction of the finite sample behaviour of the $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ tests.

(vi) For a given DGP, the one-sided t-tests have more power (in the correct tail) than the corresponding two-sided LM-type tests, as would be expected. Moreover, and consistent with both the discussion concerning theoretical power rates against local alternatives in Shao and Wu (2007a) and the simulation findings in Lobato and Robinson (1998), the KPSS-type tests have considerably lower power to detect departures from short memory than do the corresponding LM- and t-based fractional integration tests discussed in this paper, at least provided reasonable bandwidths m are chosen.

7 Empirical Examples

7.1 Bitcoin Returns

We apply the semiparametric long memory tests described in this paper to the daily returns of Bitcoin over the period 17 September 2014 to 31 December 2019, giving a total of $T = 1932$ daily observations. The data were retrieved from Yahoo Finance. The logarithm of the closing price of Bitcoin in USD is graphed in Figure 3 along with the returns series, defined as first differences of the (log) closing price series. A visual inspection of the data suggests the plausibility of changes in slope, implying changes in level at the same point in the returns series, with the most obvious case being at around the beginning of 2018. The red line on the graphs shows the fitted deterministic trend/level of the series allowing for two breaks, the locations of which are estimated by applying the RSS-based estimator discussed in Section 5 to the returns data setting $k = 2$. The estimated break dates are 24 March 2017 and 16 December 2017.

Here, and throughout the empirical examples reported in this section, we set the trimming parameters equal to the same values as were used in the Monte Carlo experiments in Section 6, that is $\tau_L = 0.15$ and $\tau_U = 0.85$. Where multiple breaks were estimated, we set the minimum spacing parameter λ defined in Section 5 to $\lambda = 0.10$.

Evidence of long memory in returns would of course be in strong violation of the efficient market hypothesis, and so it is of interest in the context of the Bitcoin return data to test $H_0 : \delta = 0$ against the alternative $H_1 : \delta > 0$. We do so using both the test based on the $t_m(0)$ statistic of Lobato and Robinson (1998), which does not allow for a level break, and the analogues of this test based on the $t_m(0; \widehat{\tau})$ and $t_m(0; \widehat{\boldsymbol{\tau}})$ statistics allowing for the presence of either one or two level breaks, respectively, in each case occurring at unknown points in the sample. Following the recommendations from our Monte Carlo study we computed the statistics for a range of values of the bandwidth parameter, m, lying between $\lfloor T^{0.5} \rfloor = 43$ and $\lfloor T^{0.65} \rfloor = 137$, inclusive, as well as for the automatic bandwidth rule, m_{LR} of Lobato and Robinson (1998) with the value that this takes reported in parentheses below the outcome of the statistics. The results are summarised in Table 1. Here, and also in Tables 3 and 4, the superscripts *, ** and *** denote outcomes which are statistically significant at the 10%, 5% and 1% level, respectively.

Using Lobato and Robinson's $t_m(0)$ test we can reject H_0 at the 10% level when using the data-dependent bandwidth rule, m_{LR}, and for all but the smallest and largest of the other bandwidths considered. The null can also be rejected at the 5% level for $m = 75$ and

Figure 3: Bitcoin daily data 9/17/2014 to 12/31/2019

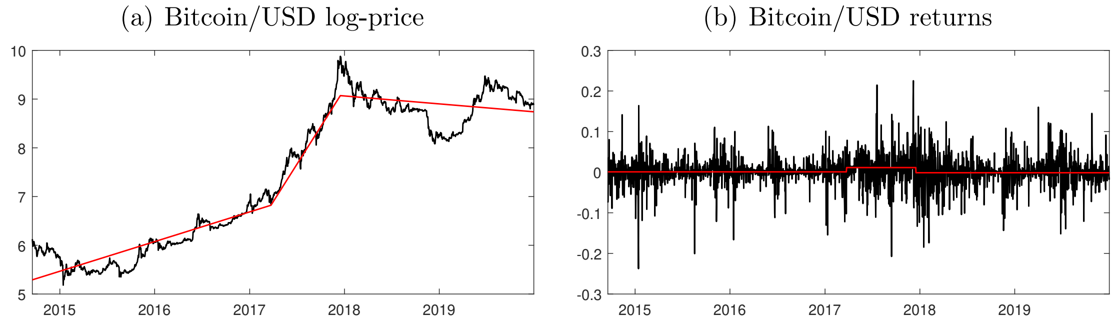


Table 1: Tests of $H_0 : \delta = 0$ versus $H_1 : \delta > 0$ in Bitcoin returns data

	m	$t_m(0)$ $k = 0$	$t_m(0; \widehat{\tau})$ $k = 1$	$t_m(0; \widehat{\boldsymbol{\tau}})$ $k = 2$
$\lfloor T^{0.50} \rfloor$	43	1.19	0.80	-0.30
	50	1.34*	0.94	-0.21
$\lfloor T^{0.55} \rfloor$	64	1.45*	1.02	-0.11
	75	1.99**	1.51*	0.31
$\lfloor T^{0.60} \rfloor$	93	2.02**	1.54*	0.33
	100	1.38*	0.91	-0.27
	125	1.57*	1.08	-0.13
$\lfloor T^{0.65} \rfloor$	137	1.09	0.60	-0.59
	m_{LR}	1.62* $(m = 510)$	1.37* $(m = 510)$	0.13 $(m = 510)$

Note: *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

$m = 93$. On balance we surmise from the results for the standard Lobato and Robinson test that the short memory null hypothesis is rejected in favour of long memory in the Bitcoin returns data. On the other hand, for the test based on $t_m(0; \widehat{\tau})$, which fits a level break to the data, the evidence against the null hypothesis is considerably weaker and, in particular, H_0 can only be rejected at the 10% level for the set of bandwidths $m \in \{75, 93, m_{LR}\}$. Allowing for two breaks, no choice of bandwidth results in a rejection at even the 10% level for the $t_m(0; \widehat{\boldsymbol{\tau}})$ test. These results suggest that the finding of long memory in Bitcoin returns by the standard Lobato and Robinson (1998) test is likely attributable to the presence of at least one level break in the returns data.

7.2 VIX Market Volatility

In the next example we consider market volatility, measured by VIX, using daily data from 1 January 2000 to 31 December 2019 for a total of $T = 5031$ observations. The data were downloaded from Yahoo Finance and are graphed in Figure 4. The red step function on the graph shows the fitted deterministic level of the series allowing for 5 level breaks.

It has been argued by several authors that long memory in volatility is an important

Figure 4: VIX daily data 1/1/2000 to 12/31/2019

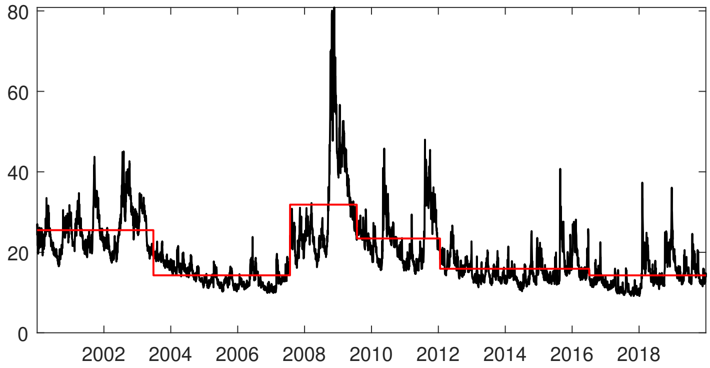


Table 2: Tests of $H_0 : \delta = 0$ versus $H_1 : \delta > 0$ in VIX volatility data

	m	$t_m(0)$ $k=0$	$t_m(0;\widehat{\tau})$ $k=1$	$t_m(0;\widehat{\boldsymbol{\tau}})$ $k=2$	$t_m(0;\widehat{\boldsymbol{\tau}})$ $k=3$	$t_m(0;\widehat{\boldsymbol{\tau}})$ $k=4$	$t_m(0;\widehat{\boldsymbol{\tau}})$ $k=5$
$\lfloor T^{0.50} \rfloor$	70	13.39	11.35	10.71	6.01	4.86	4.74
	75	14.39	12.27	11.60	6.74	5.54	5.42
	100	18.98	16.51	15.63	10.06	8.61	8.47
$\lfloor T^{0.55} \rfloor$	108	20.44	17.86	16.94	11.14	9.62	9.48
	125	23.43	20.63	19.61	13.32	11.66	11.51
	150	27.52	24.40	23.17	16.30	14.48	14.31
$\lfloor T^{0.60} \rfloor$	166	30.08	26.79	25.45	18.24	16.31	16.14
	175	31.49	28.10	26.71	19.29	17.30	17.13
	200	35.40	31.76	30.25	22.30	20.18	19.99
	225	39.17	35.30	33.68	25.23	22.96	22.76
	250	42.75	38.65	36.91	27.98	25.57	25.36
$\lfloor T^{0.65} \rfloor$	254	43.30	39.17	37.40	28.39	25.95	25.74
	m_{LR}	10.38	8.63	8.18	4.07	3.04	2.93
		$(m = 54)$	$(m = 54)$	$(m = 54)$	$(m = 54)$	$(m = 54)$	$(m = 54)$

Note: All statistics in this table are significant at the 1% level.

stylized fact; see e.g. Andersen *et al.* (2001) and references therein. Furthermore, long memory in volatility is relevant in asset pricing. For example, Baillie *et al.* (1996) use asset pricing as motivation for their FIGARCH model, and Christensen and Nielsen (2007) discuss implications of long memory in volatilty in the context of stock pricing. Other authors, however, suggest volatility might be a short memory process with the statistical evidence for long memory disappearing once level shifts in the data are accounted for; see, among others, Granger and Hyung (2004).

To investigate this further, we test the short memory null hypothesis $H_0 : \delta = 0$ against the long memory alternative $H_1 : \delta > 0$ in the VIX data. We report the outcomes of the $t_m(0)$ statistic, the $t_m(0; \widehat{\tau})$ statistic which allows for the presence of up to one level break, and the $t_m(0; \widehat{\boldsymbol{\tau}})$ statistic which allows for up k level breaks and do this for each of $k = 2, \ldots, 5$. In view of the theoretical results in Phillips (1998) we stop at $k = 5$.

We again computed these statistics for a range of values of the bandwidth parameter, m, between $\lfloor T^{0.5} \rfloor = 70$ and $\lfloor T^{0.65} \rfloor = 254$, inclusive, together with the automatic bandwidth rule, m_{LR}. The results are summarised in Table 2. Following Andersen *et al.* (2001), we also conducted the analysis using logarithmically transformed VIX data, and the results were nearly identical to those reported in Table 2.

It is seen from the results in Table 2 that the short memory null hypothesis is easily rejected at the 1% significance level for all bandwidths considered and regardless of how many level breaks we fit to the data. The results of these tests therefore strongly suggest that long memory is a feature of the VIX data, and that this would not appear to be spurious long memory due to unmodelled level breaks.

7.3 U.S. CPI Inflation

We next consider U.S. CPI inflation, defined as the first differences of the logarithm of the price index. Specifically, we used the series CPIAUCSL from the FRED database, which is the CPI for all items, Urban consumers, seasonally adjusted, base year 1984. We used monthly observations spanning January 1970 to December 2019, for $T = 599$ observations on the first differences. The log-CPI data along with the inflation data, the latter multiplied by 1200 to return a measure that is compatible with the commonly reported inflation rate, are both plotted in Figure 5. U.S. inflation is widely argued to have gone through several different policy regimes over the sample period considered here, most notably the Great Inflation period of the 1970s, the subsequent Volcker-Greenspan era of inflation rate targeting by the U.S. Federal Reserve starting in the early 1980s, and the response to the financial crisis of 2008. Figure 5 is indeed suggestive of the possibility of several level breaks in the inflation data. The red step line on the graphs again shows the fitted deterministic trend/level of the series allowing for up to four breaks. The estimated break dates are August 1977, July 1982, January 1991, and July 2008, broadly consistent with the regimes discussed above.

We again test the short memory null hypothesis, $H_0 : \delta = 0$, against the alternative of (positive) long memory in the U.S. inflation data. We consider both the test based on the $t_m(0)$ statistic of Lobato and Robinson (1998), and the corresponding tests based on the $t_m(0; \hat{\tau})$ and $t_m(0; \hat{\boldsymbol{\tau}})$ statistics allowing for the presence of up to $k = 1, \ldots, 4$ level breaks, in each case at unknown points in the sample. The results are reported in Table 3 again for a range of values of the bandwidth parameter, m, lying between $\lfloor T^{0.5} \rfloor = 24$ and $\lfloor T^{0.65} \rfloor = 63$, inclusive, and the data-dependent bandwidth rule, m_{LR}.

Lobato and Robinson's $t_m(0)$ test overwhelmingly rejects short memory at any conventional significance level for all of the bandwidths considered. Allowing for the presence of level breaks considerably reduces the magnitude of the test statistics. The test outcomes are generally still strongly significant when allowing for one or two level breaks, but when three level breaks are allowed for, the null cannot be rejected at the 5% level

Figure 5: U.S. CPI monthly data Jan. 1970 to Dec. 12/2019

(a) Log-CPI (b) Inflation

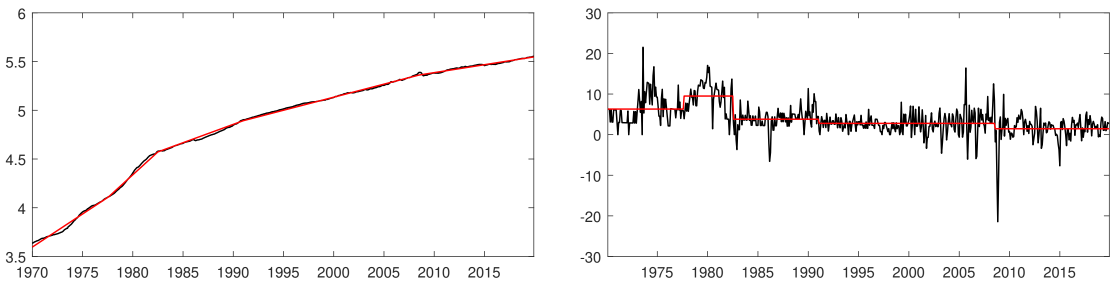


Table 3: Tests of $H_0 : \delta = 0$ versus $H_1 : \delta > 0$ in U.S. monthly CPI inflation data

	m	$t_m(0)$ $k=0$	$t_m(0;\widehat{\tau})$ $k=1$	$t_m(0;\widehat{\boldsymbol{\tau}})$ $k=2$	$t_m(0;\widehat{\boldsymbol{\tau}})$ $k=3$	$t_m(0;\widehat{\boldsymbol{\tau}})$ $k=4$
$\lfloor T^{0.50} \rfloor$	24	6.59***	2.80***	1.56*	0.50	0.15
	30	8.05***	3.55***	2.18**	1.04	0.67
$\lfloor T^{0.55} \rfloor$	33	8.76***	4.00***	2.57***	1.41*	1.02
	40	9.96***	4.32***	2.67***	1.50*	1.08
$\lfloor T^{0.60} \rfloor$	46	10.96***	4.80***	3.08***	1.89**	1.44*
	50	11.72***	5.22***	3.43***	2.22**	1.75**
	60	13.04***	5.66***	3.69***	2.41***	1.93*
$\lfloor T^{0.65} \rfloor$	63	13.40***	5.78***	3.78***	2.48***	1.98**
	m_{LR}	7.66*** ($m = 28$)	4.46*** ($m = 41$)	2.99*** ($m = 45$)	2.05** ($m = 47$)	1.62* ($m = 48$)

Note: *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

for bandwidths up to $m = 40$. When allowing for four level breaks only the tests based on bandwidths of $m = 50$ and $m = 63$ are significant at the 5% level. Consequently, while the standard Lobato and Robinson (1998) test presents very strong evidence in favour of long memory in the U.S. inflation rate, tests which allows for different policy regimes within the sample period are more suggestive that U.S. inflation is a short memory series.

7.4 Real U.S. GDP Growth Rate

Finally we consider U.S. GDP growth rates obtained as the first difference of the logarithm of real U.S. quarterly GDP (seasonally adjusted) over the period 1947Q1 to 2019Q4 obtained from the FRED database (series GDPC1), for a total of $T = 292$ quarterly observations. The data for U.S. (log) GDP and the GDP growth rates are both graphed in Figure 6. The red line on the graphs again shows the fitted deterministic trend/level of the series allowing for up to three breaks. The estimated break dates are 1973Q2, 1982Q3 and 2000Q2, broadly consistent with the first oil crisis, changes in the Fed policy (discussed in the context of the U.S. CPI data in Section 7.3) and the end of the dot-com bubble.

In particular we will test the null hypothesis that growth rates are short memory,

Figure 6: U.S. GDP quarterly data 1947Q1 to 2019Q4

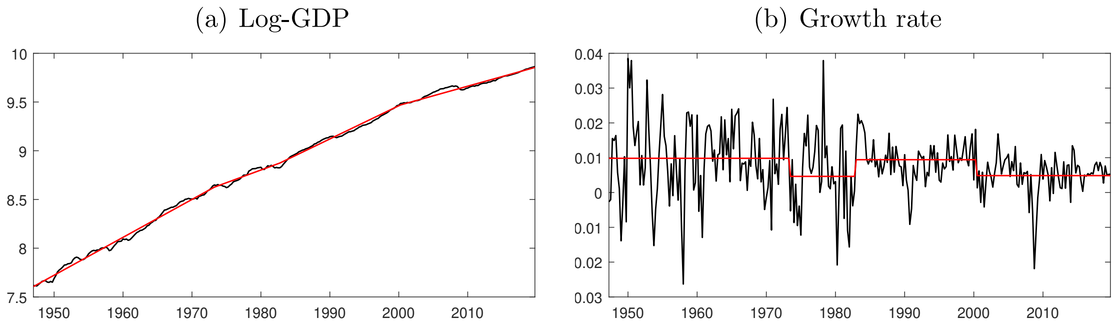


Table 4: Tests of $H_0 : \delta = 0$ versus $H_1 : \delta < 0$ in U.S. quarterly growth rates

	m	$t_m(0)$ $k=0$	$t_m(0;\widehat{\tau})$ $k=1$	$t_m(0;\widehat{\boldsymbol{\tau}})$ $k=2$	$t_m(0;\widehat{\boldsymbol{\tau}})$ $k=3$
$\lfloor T^{0.50} \rfloor$	17	-0.12	-1.07	-1.47^{*}	-1.10
	20	0.04	-0.90	-1.32^{*}	-0.89
$\lfloor T^{0.60} \rfloor$	22	0.02	-0.89	-1.30^{*}	-0.89
	25	0.30	-0.63	-1.05	-0.66
$\lfloor T^{0.60} \rfloor$	30	0.19	-0.70	-1.09	-0.73
	35	-0.52	-1.33^{*}	-1.66^{**}	-1.43^{*}
$\lfloor T^{0.65} \rfloor$	40	0.02	-0.82	-1.17	-0.93
	m_{LR}	0.10 $(m = 31)$	-0.83 $(m = 32)$	-1.70^{**} $(m = 33)$	-1.47^{*} $(m = 34)$

Note: *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

$H_0 : \delta = 0$, such that the log-level of GDP follows an $I(1)$ process, against the alternative of negative long memory (antipersistence) in growth rates, $H_1 : \delta < 0$, such that the log-level of GDP is less persistent than an $I(1)$ process. As in the previous examples, we consider the test of Lobato and Robinson (1998) based on the $t_m(0)$ statistic, and the corresponding tests based on the $t_m(0;\widehat{\tau})$ and $t_m(0;\widehat{\boldsymbol{\tau}})$ statistics allowing for up to $k = 1, 2, 3$ level breaks, in each case at unknown points in the sample. The results are reported in Table 4 again for a range of values of the bandwidth parameter, m, lying between $\lfloor T^{0.5} \rfloor = 17$ and $\lfloor T^{0.65} \rfloor = 40$, inclusive, and the data-dependent bandwidth rule, m_{LR}.

With only a few exceptions, the tests reported are unable to reject the null hypothesis that GDP growth rates are short memory against $H_1 : \delta < 0$ at conventional significance levels. The results from these tests do not therefore appear to support the conjecture of Perron (1989) that U.S. GDP is $I(0)$ about a broken linear trend, particularly when recalling that our test is of the null hypothesis that U.S. GDP is $I(1)$ around a broken trend.

8 Conclusions

We have developed semiparametric tests, based on the score and Lagrange multiplier testing principles, for the fractional order of integration of a univariate time series which may be subject to the presence of level breaks. This is of significant practical importance as it is well known that long memory and level breaks can be mistaken for one another, with unmodelled level breaks rendering standard fractional integration tests highly unreliable. Our approach generalises the tests for the null hypothesis of weak dependence ($I(0)$) developed in Lobato and Robinson (1998). These tests are based on the local Whittle approach, and therefore do not require the user to specify a parametric model for any weak autocorrelation present in the data, which is a considerable practical advantage where the confounding effects of long memory and level breaks are present. We also show how, as conjectured in Lobato and Robinson (1998, p. 478), their testing approach can be generalised to develop tests of the null hypothesis that a series is $I(\delta)$ for any δ lying in the usual stationary and invertible region of the parameter space, not just $\delta = 0$. In spite of these generalisations, our tests are shown to attain the same standard asymptotic null distributions and asymptotic local power functions as the corresponding tests in Lobato and Robinson (1998); hence, there is no loss of asymptotic local power from allowing for level breaks, even where no level breaks are present. Monte Carlo simulations suggest that the tests perform well and that the predictions from the asymptotic theory appear to hold reasonably well in finite samples. The practical relevance of our proposed tests was highlighted with a number of empirical examples relating to macroeconomics and finance.

Appendix A Mathematical Proofs

In this appendix we provide proofs of Theorems 1 and 2.

A.1 Proof of Theorem 1

We use the notation $\delta_c := cm^{-1/2}$, so that, under H_c, we have $\delta = \delta_0 + \delta_c$.

Consider first the proof under Assumption 1. We re-write $t^*_m(\delta_0)$ in (4) as

$$t^*_m(\delta_0) = \frac{-m^{-1/2}G^{-1}\sum_{j=1}^{m}\nu_j\lambda_j^{2\delta}j^{-2\delta_c}I_u(\lambda_j)}{m^{-1}G^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta}j^{-2\delta_c}I_u(\lambda_j)}. \tag{10}$$

The numerator of $t^*_m(\delta_0)$ in (10) is

$$\begin{aligned} -m^{-1/2}\sum_{j=1}^{m}\nu_j G^{-1}\lambda_j^{-2\delta_c}\lambda_j^{2\delta}I_u(\lambda_j) &= -m^{-1/2}\sum_{j=1}^{m}\nu_j j^{-2\delta_c}\left(G^{-1}\lambda_j^{2\delta}I_u(\lambda_j) - 2\pi I_\varepsilon(\lambda_j)\right) && (11)\\ &\quad - m^{-1/2}\sum_{j=1}^{m}\nu_j 2\pi I_\varepsilon(\lambda_j) && (12)\\ &\quad - m^{-1/2}\sum_{j=1}^{m}\nu_j(j^{-2\delta_c}-1)2\pi I_\varepsilon(\lambda_j). && (13)\end{aligned}$$

Letting $I_\varepsilon(\lambda_j)$ denote the periodogram of ε_t, (4.8) of Robinson (1995b) shows that, for

$r \leq m$,

$$\sum_{j=1}^{r}\left(\frac{I_u(\lambda_j)}{G\lambda_j^{-2\delta}} - 2\pi I_\varepsilon(\lambda_j)\right) = O_p\left(r^{1/3}(\ln(r))^{2/3} + r^3T^{-2} + r^{1/2}T^{-1/4}\right). \tag{14}$$

Then, letting $b_j := \nu_j j^{-2\delta_c}$ and proceeding as in Robinson (1995b) it follows that the remainder term (11) is $o_p(1)$. This involves using summation by parts, (14), and the bound $|b_j - b_{j+1}| = O(j^{-1})$, which follows by elementary calculations. From (4.11) of Robinson (1995b) it follows directly that (12) converges in distribution to $N(0,1)$.

Next, by a Taylor series expansion and by definition of δ_c, $j^{-2\delta_c} = 1 - 2cm^{-1/2}(\ln j) + 2c^2m^{-1}(\ln j)^2 j^{-2\delta_{mvt}}$ for $|\delta_{mvt}| \leq |\delta_c|$, so that (13) is

$$2cm^{-1}\sum_{j=1}^{m}\nu_j(\ln j)2\pi I_\varepsilon(\lambda_j) + 2c^2m^{-3/2}\sum_{j=1}^{m}\nu_j(\ln j)^2 j^{-2\delta_{mvt}}2\pi I_\varepsilon(\lambda_j). \tag{15}$$

Writing $\ln j = \nu_j + m^{-1}\sum_{k=1}^{m}\ln k$, the first term of (15) is

$$2cm^{-1}\sum_{j=1}^{m}\nu_j^2 2\pi I_\varepsilon(\lambda_j) + 2cm^{-1}\left(m^{-1}\sum_{k=1}^{m}\ln k\right)\sum_{j=1}^{m}\nu_j 2\pi I_\varepsilon(\lambda_j).$$

Noting that $2cm^{-1}\sum_{j=1}^{m}\nu_j^2 2\pi E(I_\varepsilon(\lambda_j)) = 2cm^{-1}\sum_{j=1}^{m}\nu_j^2 \to 2c$, the first term converges in probability to $2c$ by a law of large numbers. Using the result for (12) and the fact that $m^{-1}\sum_{k=1}^{m}\ln k = O(\ln m)$, the second term is $O_p(m^{-1/2}\ln m) = o_p(1)$. Next, the expectation of the absolute value of the second term of (15) is

$$O\left(m^{-3/2}\sum_{j=1}^{m}\nu_j(\ln j)^2 j^{2|\delta_c|}\right) = O\left(m^{-1/2}(\ln m)^3 m^{2|\delta_c|}\right) = o(1),$$

where the last equality follows because $m^{1/2} \geq \ln m$, which implies $m^{2|\delta_c|} \leq m^{2|c|/\ln m} = e^{2|c|}$. This shows that the second term of (15) converges to zero in L_1-norm and hence in probability.

The denominator of $t_m^*(\delta_0)$ in (10) may be analyzed in the same way to establish the result that $m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta}j^{-2\delta_c}2\pi I_u(\lambda_j) \to_p G$. The claim of Theorem 1 under Assumption 1 follows by combining these results.

Next, we prove the theorem under Assumption 2. Instead of the bound (14) from (4.8) of Robinson (1995b), we let $\alpha_T(\lambda) := (1 - e^{i\lambda})^{-(\delta_0+\delta_c)}$ and use Lemma 4 of Shao and Wu (2007a), where it is shown that, under Assumption 2,

$$\sum_{j=1}^{r}\left(\frac{I_u(\lambda_j)}{|\alpha_T(\lambda_j)|^2 f(\lambda_j)} - \frac{I_\eta(\lambda_j)}{f(\lambda_j)}\right) = O_p\left(r^{1/4}(\ln r)^{1/2} + r^{1/2}T^{-1/4}\right). \tag{16}$$

Denoting $\alpha(\lambda_j) := (1 - e^{i\lambda})^{-\delta_0}$ and $\alpha_c(\lambda_j) := (1 - e^{i\lambda})^{-\delta_c}$, so that $\alpha_T(\lambda_j) = \alpha(\lambda_j)\alpha_c(\lambda_j)$, the (scaled negative) numerator of $t_m^*(\delta_0)$ in (4) is

$$
\begin{aligned}
G^{-1}\sum_{j=1}^{m}\nu_j\lambda_j^{2\delta_0}I_u(\lambda_j) &= \sum_{j=1}^{m}\nu_j G^{-1}\lambda_j^{-2\delta_c}\lambda_j^{2\delta}I_u(\lambda_j)\\
&= \sum_{j=1}^{m}\nu_j G^{-1}\lambda_j^{-2\delta_c}|\alpha_c(\lambda_j)|^{-2}f(\lambda_j)^{-1}|\alpha_c(\lambda_j)|^2 f(\lambda_j)|\alpha_T(\lambda_j)|^2|\alpha_T(\lambda_j)|^{-2}\lambda_j^{2\delta}I_u(\lambda_j)\\
&= \sum_{j=1}^{m}\nu_j|\alpha_c(\lambda_j)|^2|\alpha_T(\lambda_j)|^{-2}f(\lambda_j)^{-1}I_u(\lambda_j)+o_p(m^{1/2}), \qquad (17)
\end{aligned}
$$

where the last equality follows by using bounds for the low-frequency approximation of the ratio of $f(\lambda_j)$ to G, see Assumption 2(iii), and of $|\alpha_T(\lambda_j)|^2$ to $\lambda_j^{-2\delta}$ as in Robinson (1995b).

For the leading term in (17), we let $b_j := \nu_j|\alpha_c(\lambda_j)|^2$ (with slight abuse of notation), and re-write it as

$$
\sum_{j=1}^{m} b_j\left(|\alpha_T(\lambda_j)|^{-2}f(\lambda_j)^{-1}I_u(\lambda_j)-f(\lambda_j)^{-1}I_\eta(\lambda_j)\right) \qquad (18)
$$

$$
+\sum_{j=1}^{m} b_j f(\lambda_j)^{-1}I_\eta(\lambda_j). \qquad (19)
$$

As in the analysis of (11) it holds that (18) is $o_p(1)$ using (16). The term (19) is asymptotically normal as shown in Shao and Wu (2007a). As in the previous case, the same arguments also give $m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta}2\pi I_u(\lambda_j)\to_p G$, and the claim of Theorem 1 under Assumption 2 follows combining these results.

A.2 Proof of Theorem 2

Let $W(\tau;d)$ denote a type I fractional Brownian motion; that is, with $E(W(\tau;d)^2)=\tau^{2(d-1/2)}$, and $W(\tau;0)$ a standard brownian motion. Also let $\kappa(d):=\frac{A(d)}{\Gamma(d+1)}$, where $A(d):=\left(\frac{1}{1+2\delta}+\int_0^\infty((1+s)^d-s^d)^2ds\right)^{1/2}$. Then we have the following result.

Lemma 1. *Under H_c and either Assumption 3 or Assumption 4, $T^{-(1/2+\delta_0)}\sum_{t=1}^{\lfloor\tau T\rfloor}u_t\Rightarrow\kappa(\delta_0)(2\pi f(0))^{1/2}W(\tau;\delta_0)$ as a càdlàg process indexed by $\tau\in\Lambda$.*

Proof of Lemma 1. Given that $\delta_0<1/2$, for m large enough there is $\bar\delta=1/2-\epsilon$ such that $\delta_0<\bar\delta$ and $\delta<\bar\delta$. Using a mean value theorem expansion,

$$
\begin{aligned}
u_t &= \Delta^{-\delta}\eta_t=\Delta^{-\delta_0}\eta_t+(cm^{-1/2})(-\ln\Delta)\Delta^{-\delta_0}\eta_t\\
&\quad+1/2(cm^{-1/2})^2(-\ln\Delta)^2\Delta^{-\delta_0}\eta_t+\ldots+1/(k!)(cm^{-1/2})^k(-\ln\Delta)^k\Delta^{-\delta_{mvt}}\eta_t, \qquad (20)
\end{aligned}
$$

where $|\delta_{mvt}-\delta_0|<|\delta-\delta_0|$ and k is an integer to be chosen.

From the fractional FCLT, the first term on the right-hand side of (20) satisfies

$$
T^{-(1/2+\delta_0)}\sum_{t=1}^{\lfloor\tau T\rfloor}\Delta^{-\delta_0}\eta_t\Rightarrow\kappa(\delta_0)(2\pi f(0))^{1/2}W(\tau;\delta_0)
$$

in the Skorohod metric; see, e.g., Hosoya (2005) and Wu and Shao (2006). Moreover, because the jumps in the partial sums take place at fixed points in time, and the limit $W(\tau;\delta)$ is a.s. continuous, the weak convergence also takes place in the uniform metric.

By the same argument, including a slowly varying function, it follows that

$$\frac{1}{\ln(T)}T^{-(1/2+\delta_0)}\sum_{t=1}^{\lfloor\tau T\rfloor}(-\ln\Delta)\Delta^{-\delta_0}\eta_t = O_p(1),$$

and hence $cm^{-1/2}T^{-(1/2+\delta_0)}\sum_{t=1}^{\lfloor\tau T\rfloor}(-\ln\Delta)\Delta^{-\delta_0}\eta_t = o_p(1)$; in both cases uniformly in τ. The $k-2$ remaining terms in the expansion of $\Delta^{-\delta}\eta_t$ in (20) can be analyzed the same way.

For the last term on the right-hand side of (20), notice that

$$\begin{aligned} E((-\ln\Delta)^k\Delta^{-\delta_{mvt}}\eta_t)^2 &= \int_{-\pi}^{\pi}|1-e^{i\lambda}|^{-2\delta_{mvt}}(\ln|1-e^{i\lambda}|)^{2k}f(\lambda)d\lambda \\ &\leq \int_{-\pi}^{\pi}|1-e^{i\lambda}|^{-2\bar{\delta}}(\ln|1-e^{i\lambda}|)^{2k}f(\lambda)d\lambda < \infty, \end{aligned}$$

where we recall that $f(\lambda)$ is the spectral density of η_t, which is bounded, uniformly in λ, under either Assumption 3 or Assumption 4. Then, by the Cauchy-Schwarz inequality,

$$\sum_{t=1}^{\lfloor\tau T\rfloor}(-\ln\Delta)^k\Delta^{-\delta_{mvt}}\eta_t \leq \left(\sum_{t=1}^{T}((-\ln\Delta)^k\Delta^{-\delta_{mvt}}\eta_t)^2\right)^{1/2}T^{1/2} = O_p(T),$$

and note that this is uniform in τ. So, upon choosing k finite but sufficiently large, $T^{-(1/2+\delta_0)}m^{-k/2}T \to 0$ by Assumption 3(iv) or Assumption 4(iv), and consequently

$$T^{-(1/2+\delta_0)}\frac{(cm^{-1/2})^k}{k!}\sum_{t=1}^{\lfloor\tau T\rfloor}(\ln\Delta)^k\Delta^{-\delta_{mvt}}\eta_t = o_p(1).$$

Combining these arguments we obtain the desired result. □

In what follows, results for stochastic of functionals of τ are to be considered as uniform in τ, unless otherwise specified. We omit the the reference to uniformity in τ for brevity.

A.2.1 Proof of Theorem 2(i)

We prove that, when $\beta_2 = 0$, $t_m(\delta_0;\tau) - t_m^*(\delta_0) = o_p(1)$ uniformly in τ. It is sufficient to show that

$$m^{-1/2}\sum_{j=1}^{m}\nu_j\lambda_j^{2\delta_0}I_{\hat{u}(\tau)}(\lambda_j) - m^{-1/2}\sum_{j=1}^{m}\nu_j\lambda_j^{2\delta_0}I_u(\lambda_j) = o_p(1), \tag{21}$$

$$m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta_0}I_{\hat{u}(\tau)}(\lambda_j) - m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta_0}I_u(\lambda_j) = o_p(1). \tag{22}$$

We give only the proof of (21). The proof of (22) is almost identical leaving out the factor ν_j and noting the different normalization.

We first note that

$$w_{\widehat{u}(\tau)}(\lambda_j) = (\beta - \widehat{\beta}(\tau))' w_{x(\tau)}(\lambda_j) + w_u(\lambda_j) = -\widehat{\beta}_2(\tau) w_{DU(\tau)}(\lambda_j) + w_u(\lambda_j),$$
$$I_{\widehat{u}(\tau)}(\lambda_j) = \widehat{\beta}_2(\tau)^2 I_{DU(\tau)}(\lambda_j) + I_u(\lambda_j) - 2\,\mathrm{Re}\left(w_{DU(\tau)}(\lambda_j) w_u(-\lambda_j)\right)\widehat{\beta}_2(\tau).$$

The absolute value of the left-hand side of (21) is then

$$\left| m^{-1/2} \sum_{j=1}^{m} \nu_j \lambda_j^{2\delta_0} I_{DU(\tau)}(\lambda_j) \widehat{\beta}_2(\tau)^2 - 2m^{-1/2} \sum_{j=1}^{m} \nu_j \lambda_j^{2\delta_0}\, \mathrm{Re}\left(w_{DU(\tau)}(\lambda_j) w_u(-\lambda_j)\right) \widehat{\beta}_2(\tau) \right|$$
$$\leq m^{-1/2} \sum_{j=1}^{m} |\nu_j| \lambda_j^{2\delta_0} I_{DU(\tau)}(\lambda_j) \widehat{\beta}_2(\tau)^2 \tag{23}$$
$$+ Cm^{-1/2} \sum_{j=1}^{m} |\nu_j| \lambda_j^{2\delta_0} \left|\mathrm{Re}\left(w_{DU(\tau)}(\lambda_j) w_u(-\lambda_j)\right)\right| |\widehat{\beta}_2(\tau)|. \tag{24}$$

From Iacone (2010) we have the bound

$$I_{DU(\tau)}(\lambda_j) \leq C(j/T)^{-1} j^{-1}, \tag{25}$$

and the bound $\widehat{\beta}_2(\tau) = O_p(T^{-1/2+\delta_0})$ follows from an application of Lemma 1 to the regression estimate. Applying also the simple bound $|\nu_j| = O(\ln m)$, the term (23) is

$$m^{-1/2} \sum_{j=1}^{m} |\nu_j| \lambda_j^{2\delta_0} I_{DU(\tau)}(\lambda_j) \widehat{\beta}_2(\tau)^2 = O_p\left((\ln m) m^{-1/2} \sum_{j=1}^{m} \left(\frac{j}{T}\right)^{2\delta_0} \left(\frac{j}{T}\right)^{-1} j^{-1} T^{-1+2\delta_0} \right)$$
$$= O_p\left((\ln m) m^{-1/2} \sum_{j=1}^{m} j^{2\delta_0 - 2} \right) = O_p\left((\ln m) m^{-1/2}\right) = o_p(1).$$

For the term (24), first note that $E\left|\lambda_j^{2\delta} I_u(\lambda_j)\right| < C$, see Lemma 3 of Shao and Wu (2007a), the proof of which also applies under Assumption 3. Thus, (24) is bounded by

$$Cm^{-1/2} \sum_{j=1}^{m} |\nu_j| \lambda_j^{2\delta_0} \left|\mathrm{Re}\left(w_{DU(\tau)}(\lambda_j) w_u(-\lambda_j)\right)\right| |\widehat{\beta}_2(\tau)|$$
$$= O_p\left((\ln m) m^{-1/2} \sum_{j=1}^{m} \left(\frac{j}{T}\right)^{2\delta_0} \left(\frac{j}{T}\right)^{-1/2} j^{-1/2} \left(\frac{j}{T}\right)^{-\delta} T^{-1/2+\delta_0} \right)$$
$$= O_p\left((\ln m) m^{-1/2} T^{\delta-\delta_0} \sum_{j=1}^{m} j^{(\delta_0 - 1) + (\delta_0 - \delta)} \right).$$

Note that, by mean value expansion, $T^{\delta-\delta_0} = 1 + cm^{(-1/2)}(\ln T) T^{\delta_{mvt}-\delta_0}$, where $|\delta_{mvt} - \delta_0| \leq |\delta - \delta_0|$. Thus, for T large enough, $|T^{\delta_0-\delta} - 1| < cm^{(-1/2)}(\ln T) T^{\epsilon}$ for any $\epsilon > 0$ arbitrarily small. Thus, $T^{\delta_0-\delta} \to 1 = O(1)$. The term $j^{(\delta_0-1)+(\delta_0-\delta)}$ can be discussed in the same way. Therefore, the stochastic order of (24) is

$$O_p\left((\ln m)m^{-1/2}\sum_{j=1}^{m} j^{(\delta_0-1)}\right) = O_p\left((\ln m)^2 m^{-1/2} m^{\max(\delta_0,0)}\right),$$

which is $o_p(1)$ recalling that $\delta_0 < 1/2$.

A.2.2 Proof of Theorem 2(ii)

We prove that, when $\beta_2 \neq 0$, $t_m(\delta_0;\widehat{\tau}) - t^*_m(\delta_0) = o_p(1)$. Because $\beta_2 \neq 0$, we have to account for the difference between $\widehat{\tau}$ and τ^*. We first notice that Assumption H1 of Lavielle and Moulines (2000) holds both under Assumption 1 and under Assumption 2, in view of the stationarity of u_t. Following Theorem 7 of Lavielle and Moulines (2000), for any $|\delta| < 1/2$,

$$\widehat{\tau} - \tau^* = O_p(T^{-1}). \tag{26}$$

Proceeding as Bai (1994) and Iacone *et al.* (2019), we also establish $\beta_2 - \widehat{\beta}_2(\widehat{\tau}) = O_p(T^{\delta_0 - 1/2})$. Notice here that the key step of the proof in Iacone *et al.* (2019) exploits the rate in (26), the Hájek-Rényi-type inequality in (8) of Lavielle and Moulines (2000), and the fractional FCLT.

We now re-write

$$\begin{aligned}\widehat{u}_t(\widehat{\tau}) &= \beta_1 + \beta_2 DU_t(\tau^*) + u_t - \widehat{\beta}_1(\widehat{\tau}) - \widehat{\beta}_2(\widehat{\tau}) DU_t(\widehat{\tau}) \\ &= \beta_1 + \beta_2 DU_t(\tau^*) - \widehat{\beta}_2(\widehat{\tau}) DU_t(\tau^*) + \widehat{\beta}_2(\widehat{\tau}) DU_t(\tau^*) + u_t - \widehat{\beta}_1(\widehat{\tau}) - \widehat{\beta}_2(\widehat{\tau}) DU_t(\widehat{\tau}) \\ &= (\beta_1 - \widehat{\beta}_1(\widehat{\tau})) + (\beta_2 - \widehat{\beta}_2(\widehat{\tau})) DU_t(\tau^*) + u_t + \widehat{\beta}_2(\widehat{\tau})(DU_t(\tau^*) - DU_t(\widehat{\tau})),\end{aligned}$$

and thus

$$\begin{aligned}I_{\widehat{u}(\widehat{\tau})}(\lambda_j) = I_u(\lambda_j) &+ (\beta_2 - \widehat{\beta}_2(\widehat{\tau}))^2 I_{DU(\tau^*)}(\lambda_j) + \widehat{\beta}_2(\widehat{\tau})^2 I_{(DU(\tau^*)-DU(\widehat{\tau}))}(\lambda_j) \\ &+ 2\,\mathrm{Re}\left(w_{DU(\tau^*)}(\lambda_j) w_u(-\lambda_j)\right)(\beta_2 - \widehat{\beta}_2(\widehat{\tau})) \\ &+ 2(\beta_2 - \widehat{\beta}_2(\widehat{\tau}))\,\mathrm{Re}\left(w_{DU(\tau^*)}(\lambda_j) w_{DU_t(\tau^*)-DU_t(\widehat{\tau})}(-\lambda_j)\right)\widehat{\beta}_2(\widehat{\tau}) \\ &+ 2\,\mathrm{Re}\left(w_u(\lambda_j) w_{DU_t(\tau^*)-DU_t(\widehat{\tau})}(-\lambda_j)\right)\widehat{\beta}_2(\widehat{\tau}).\end{aligned}$$

Proceeding as in the proof of part (i), we need to show that

$$m^{-1/2}(\ln m)\sum_{j=1}^{m}\lambda_j^{2\delta_0}\left|(\beta_2 - \widehat{\beta}_2(\widehat{\tau}))^2 I_{DU(\tau^*)}(\lambda_j)\right| = o_p(1), \tag{27}$$

$$m^{-1/2}(\ln m)\sum_{j=1}^{m}\lambda_j^{2\delta_0}\left|2\,\mathrm{Re}\left(w_{DU(\tau^*)}(\lambda_j) w_u(-\lambda_j)\right)(\beta_2 - \widehat{\beta}_2(\widehat{\tau}))\right| = o_p(1), \tag{28}$$

and

$$m^{-1/2}(\ln m)\left|\sum_{j=1}^{m}\widehat{\beta}_2(\widehat{\tau})\lambda_j^{2\delta_0}I_{(DU(\tau^*)-DU(\widehat{\tau}))}(\lambda_j)\right| = o_p(1), \tag{29}$$

$$m^{-1/2}(\ln m)\sum_{j=1}^{m}|\beta_2-\widehat{\beta}_2(\widehat{\tau})|\lambda_j^{2\delta_0}\left|\text{Re}\left(w_{DU(\tau^*)}(\lambda_j)w_{DU_t(\tau^*)-DU_t(\widehat{\tau})}(-\lambda_j)\right)\right||\widehat{\beta}_2(\widehat{\tau})| = o_p(1), \tag{30}$$

$$m^{-1/2}(\ln m)\sum_{j=1}^{m}\lambda_j^{2\delta_0}\left|\text{Re}\left(w_u(\lambda_j)w_{DU_t(\tau^*)-DU_t(\widehat{\tau})}(-\lambda_j)\right)\right||\widehat{\beta}_2(\widehat{\tau})| = o_p(1). \tag{31}$$

The bounds in (27) and in in (28) follow as in the discussion of (23) and (24), respectively.

Next, assuming without loss of generality that $\widehat{\tau} < \tau^*$,

$$w_{(DU(\tau^*)-DU(\widehat{\tau}))}(\lambda) = \frac{1}{\sqrt{2\pi T}}\sum_{t=\lfloor\widehat{\tau}T\rfloor}^{\lfloor\tau^*T\rfloor+1}e^{i\lambda},$$

$$\left|w_{(DU(\tau^*)-DU(\widehat{\tau}))}(\lambda)\right| \le \frac{1}{\sqrt{2\pi T}}(\lfloor\tau^*T\rfloor-\lfloor\widehat{\tau}T\rfloor) = O_p(T^{-1/2}),$$

and hence $I_{DU(\tau^*)-DU(\widehat{\tau})}(\lambda_j) = O_p(T^{-1})$. Thus, since $\widehat{\beta}_2(\widehat{\tau}) \to_p \beta_2$, (29) is of order

$$O_p\left(m^{-1/2}(\ln m)\sum_{j=1}^{m}j^{2\delta_0}T^{-2\delta_0}T^{-1}\right) = O_p\left(m^{-1/2}(\ln m)T^{-2\delta_0-1}m^{2\delta_0+1}\right) = o_p(1).$$

The term (30) is, using (25),

$$O_p\left(m^{-1/2}(\ln m)T^{\delta_0-1/2}\sum_{j=1}^{m}j^{2\delta_0}T^{-2\delta_0}(T^{1/2}j^{-1}T^{-1/2})\right)$$
$$= O_p\left(m^{-1/2}(\ln m)T^{-\delta_0-1/2}\sum_{j=1}^{m}j^{2\delta_0-1}\right) = O_p\left(m^{-1/2}(\ln m)^2T^{-\delta_0-1/2}m^{\max(2\delta_0,0)}\right),$$

which is $o_p(1)$.

Finally, (31) is of order

$$O_p\left(m^{-1/2}(\ln m)\sum_{j=1}^{m}j^{\delta_0+(\delta_0-\delta)}T^{-\delta_0-(\delta_0-\delta)}T^{-1/2}\right) = O_p\left(m^{\delta_0+1/2}(\ln m)T^{-\delta_0-1/2}\right) = o_p(1).$$

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics* 61, 43–76.

Bai, J. (1994). Least squares estimation of a shift in linear processes. *Journal of Time*

Series Analysis 15, 453–472.

Baillie, R.T., Bollerslev, T., and Mikkelsen H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 3–30.

Busetti, F., and Harvey, A. (2001). Testing for the presence of a random walk in series with structural breaks. *Journal of Time Series Analysis* 22, 127–150.

Busetti, F., and Harvey, A. (2003). Further comments on stationarity tests in series with structural breaks at unknown points. *Journal of Time Series Analysis* 24, 137–140.

Christensen, B.J., and Nielsen, M.Ø. (2007). Effect of long memory in volatility on stock market fluctuations. *The Review of Economics and Statistics* 89, 684–700.

Diebold, F.X., and Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics* 105, 131–159.

Geweke, J., and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4, 221–238.

Granger, C.W.J., and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the S&P500 absolute stock returns. *Journal of Empirical Finance* 11, 399–421.

Hosoya, Y. (2005). Fractional invariance principle. *Journal of Time Series Analysis* 26, 463–486.

Iacone, F. (2010). Local Whittle estimation of the memory parameter in presence of deterministic components. *Journal of Time Series Analysis* 31, 37–49.

Iacone, F., Leybourne, S.J., and Taylor, A.M.R. (2013). Testing for a break in trend when the order of integration is unknown. *Journal of Econometrics* 176, 30–45.

Iacone, F., Leybourne, S.J., and Taylor, A.M.R. (2019). Testing the order of fractional integration of a time series in the possible presence of a trend break at an unknown point. *Econometric Theory* 35, 1201–1233.

Johansen, S., and Nielsen, M.Ø. (2012). A necessary moment condition for the fractional functional central limit theorem. *Econometric Theory* 28, 671–679.

Künsch, H. (1987). Statistical aspects of self-similar processes. *Proceedings of the First World Congress of the Bernoulli Society*, vol. 1, 67–74.

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. *Journal of Econometrics* 54, 159–178.

Lavielle, M., and Moulines, E. (2000). Least-squares estimation of an unknown number of shifts in a time series. *Journal of Time Series Analysis* 21, 33–59.

Lobato, I.N., and Robinson, P.M. (1998). A nonparametric test for I(0). *Review of Economic Studies* 65, 475–495.

Nielsen, M.Ø. (2004). Effcient likelihood inference in nonstationary univariate models.

Econometric Theory 20, 116–146.

Perron, P. (1989). The Great Crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57, 1361–401.

Phillips, P.C.B. (1998). New tools for understanding spurious regressions. *Econometrica* 66, 1299–1325.

Phillips, P.C.B., and Shimotsu, K. (2004). Local Whittle estimation in nonstationary and unit root cases. *Annals of Statistics* 32, 656–692.

Robinson, P.M. (1994). Effcient tests of nonstationary hypotheses. *Journal of the American Statistical Association* 89, 1420–1437.

Robinson, P.M. (1995a). Log-periodogram regression of time series with long range dependence. *Annals of Statistics* 23, 1048–1072.

Robinson, P.M. (1995b). Gaussian semiparametric estimation of long range dependence. *Annals of Statistics* 23, 1630–1661.

Shao, X., and Wu, W.B. (2007a). Local asymptotic powers of nonparametric and semiparametric tests for fractional integration. *Stochastic Processes and their Applications* 117, 251–261.

Shao, X., and Wu, W.B. (2007b). Local Whittle estimation of fractional integration for nonlinear processes. *Econometric Theory* 23, 899–929.

Shimotsu, K., and Phillips, P.C.B. (2006). Local Whittle estimation of fractional integration and some of its variants. *Journal of Econometrics* 130, 209–233.

Tanaka, K. (1999). The nonstationary fractional unit root. *Econometric Theory* 15, 549–582.

Wu, W.B., and Shao, X. (2006). Invariance principles for fractionally integrated nonlinear processes. *IMS Lecture Notes—Monograph Series: Recent Developments in Nonparametric Inference and Probability* 50, 20–30.

Supplementary Appendix to

“Semiparametric Tests for the Order of Integration in the Possible Presence of Level Breaks”

by

Fabrizio Iacone
Università degli Studi di Milano
and University of York

Morten Ørregaard Nielsen
Queen's University
and CREATES

A.M. Robert Taylor*
University of Essex

June 15, 2020

Abstract

This supplementary appendix contains a Monte Carlo simulation study investigating the finite sample performance of the tests discussed in the main paper. Section S.1 provides details of the Monte Carlo simulation designs considered. Sections S.2 and S.3 provide details of the DGPs used to investigate empirical size and power, respectively. All results are discussed in Section 6 in the main paper.

Equation references (S.n) for $n \geq 1$ refer to equations in this supplementary appendix and other equation references are to the main paper. Additional bibliographic references are included at the end of the supplement.

S.1 Monte Carlo Design

We conducted a Monte Carlo simulation study to investigate the finite sample performance of our proposed tests based on the $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ statistics, exploring cases where no trend break occurs and where a trend break occurs. We investigate both finite sample size under the null hypothesis and finite sample power. As benchmarks for comparison, we also simulate the tests of Lobato and Robinson (1998), based on the $LM_m(\delta_0)$ and $t_m(\delta_0)$ statistics, neither of which allows for the presence of a level break; recall that, strictly speaking, Lobato and Robinson (1998) only considered tests of the short memory null hypothesis, $H_0 : \delta = 0$, based on the $LM_m(0)$ and $t_m(0)$ statistics. The behaviour of these tests in the no break case where $\beta_2 = 0$ may be viewed as

*Correspondence to: Robert Taylor, Essex Business School, University of Essex, Colchester, CO4 3SQ, U.K. E-mail: robert.taylor@essex.ac.uk

important benchmarks to compare the $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ tests against. Notice, moreover, that the results given for $LM_m(\delta_0)$ and $t_m(\delta_0)$ for $\beta_2 = 0$ coincide with the results that would obtain for the infeasible $LM^*_m(\delta_0)$ and $t^*_m(\delta_0)$ tests for *any* value of β_2 and so the results given for $LM_m(\delta_0)$ and $t_m(\delta_0)$ for $\beta_2 = 0$ also provide a useful benchmark to compare the $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ tests against for cases where $\beta_2 \neq 0$. As further benchmarks, we also include the intermediate, and infeasible, tests based on the $LM_m(\delta_0; \tau^*)$ and $t_m(\delta_0; \tau^*)$ statistics; these statistics assume knowledge of the location of the level break, τ^*, but not of its magnitude, β_2. For tests based on the $t_m(\delta_0)$, $t_m(\delta_0; \tau^*)$ and $t_m(\delta_0; \widehat{\tau})$ statistics, results are reported for both left- and right-tailed tests.

Comparison is also made with the stationarity test of KPSS, which does not allow for the presence of a level break. Although designed for testing the null hypothesis of short memory against a unit root alternative, as discussed in the introduction this test does have power against positive long memory alternatives. To account for a level break ($\beta_2 \neq 0$), Busetti and Harvey (2001, 2003) recommend computing the $KPSS$ statistic using the OLS residuals obtained from a regression on $(1, DU_t(\tau^*))'$ or $(1, DU_t(\widehat{\tau}))'$; we also include these tests in our comparisons, denoting them by $KPSS(\tau^*)$ and $KPSS(\widehat{\tau})$, respectively. Notice that, in contrast to what happens in the case of the LM-type statistics discussed in this paper, the additional regressor $DU_t(\tau^*)$, alters the limiting null distribution and, hence, the critical value of the KPSS tests.

All reported experiments are run over 10,000 Monte Carlo replications using the `RNDN` function of Gauss 20. Our simulation DGP is given by (1) with $\beta_1 = 0$ (without loss of generality) and $\beta_2 \in \{0, 0.5, 1, 2\}$, with the break location set as $\tau^* = 0.5$. Notice that $LM_m(\delta_0; \tau^*)$, $t_m(\delta_0; \tau^*)$ and $KPSS(\tau^*)$ are exact invariant with respect to β_2. For those tests involving the RSS-based break location estimate, $\widehat{\tau}$, we set $\Lambda = [0.15, 0.85]$. All reported results relate to a nominal asymptotic 0.05 level using the relevant critical value from the χ^2_1 or $N(0,1)$ distribution as appropriate. Results are reported for samples of size $T = 512$ and $T = 1024$.

To make the LM-type tests discussed in this paper operational we need to specify a bandwidth, m. The finite sample behaviour of these tests is strongly dependent on the choice of m, as shown in Lobato and Robinson (1998) in the context of the $LM_m(0)$ and $t_m(0)$ tests. Like Lobato and Robinson (1998) we will explore the behaviour of our proposed tests for both a range of fixed bandwidths and for a data-driven rule. For the former we will consider $m = \lfloor T^a \rfloor$ for $a \in \{0.40, 0.45, \ldots, 0.80\}$. For the latter we will consider an automatic bandwidth of the type developed by Lobato and Robinson (1998) in the context of the $LM_m(0)$ and $t_m(0)$ tests. This is based on the assumption that η_t follows an AR(1) process, and is given by

$$m_{LR} := \left(\frac{3T}{4\pi}\right)^{4/5} \left| \frac{-\widetilde{\phi}}{(1-\widetilde{\phi})^2} \right|^{-2/5}, \tag{S.1}$$

where $\widetilde{\phi}$ is the first-order sample autocorrelation of y_t, with m_{LR} set equal to $\lfloor 0.06T^{0.8} \rfloor$ if it falls below this lower bound and equal to $\lfloor 1.2T^{0.8} \rfloor$ if it lies above this upper bound.[1]

The m_{LR} automatic bandwidth in (S.1) does not account for the possibility of a level break in (1), as would be needed for use with the $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ tests and the $LM_m(\delta_0; \tau^*)$ and $t_m(\delta_0; \tau^*)$ tests. This can be accommodated simply by estimating $\widetilde{\phi}$ from the residuals of the regression of y_t on $(1, DU_t(\tau^*))'$ for the $LM_m(\delta_0; \tau^*)$ and $t_m(\delta_0; \tau^*)$ tests, or on $(1, DU_t(\widehat{\tau}))'$ for the $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ tests.

Furthermore, the m_{LR} rule in (S.1) is designed for tests of the null hypothesis $H_0 : \delta = 0$. For the generic null hypothesis, $H_0 : \delta = \delta_0$, it needs to be modified as follows. Introduce the operator $\Delta_+^{\delta_0}$, such that for a generic series ζ_t, $\Delta_+^{\delta_0}\zeta_t := \sum_{s=1}^{t} \pi_{t-s}\zeta_s$, where π_t denotes the term t of the binomial expansion of Δ^{δ_0}. The m_{LR} automatic bandwidth for generic δ_0 is then given by (S.1), where $\widetilde{\phi}$ is the first-order autocorrelation of $\Delta_+^{\delta_0}\xi_t$ and ξ_t denotes the residuals from a regression of y_t on 1 for the $LM_m(\delta_0)$ and $t_m(\delta_0)$ tests, on $(1, DU_t(\tau^*))'$ for the $LM_m(\delta_0; \tau^*)$ and $t_m(\delta_0; \tau^*)$ tests, or on $(1, DU_t(\widehat{\tau}))'$ for the $LM_m(\delta_0; \widehat{\tau})$ and $t_m(\delta_0; \widehat{\tau})$ tests.

The KPSS-type tests require an estimate of the long-run variance of η_t. As is standard, we adopt a sums-of-covariances based estimator using the Bartlett kernel with bandwidth M. Regarding the latter, we follow Lobato and Robinson (1998) and use

$$M_{LR} := \min\left(\left\lfloor 1.1447 \left(4\widetilde{\phi}^2 T (1-\widetilde{\phi})^{-2} (1+\widetilde{\phi})^{-2} \right)^{1/3} \right\rfloor, \left\lfloor 12\, (T/100)^{1/5} \right\rfloor \right),$$

where $\widetilde{\phi}$ is defined in the same way as for the corresponding m_{LR} bandwidth. We also report results for the deterministic bandwidth choice, $M_e := \lfloor 4(T/100)^{2/9} \rfloor$, which is the default setting in EViews if the automatic bandwidth option is not used; see EViews 10 User's Guide 2, Equation (39.43). This is also the default starting setting in the automatic bandwidth selection in Newey and West (1994).

S.2 Simulation DGPs: Empirical Size

We investigate the empirical size properties of the tests, generating u_t in (1) as an ARFIMA(1,δ,0) process; that is, $u_t = \Delta^{-\delta}\eta_t$, $\eta_t = \phi\eta_{t-1} + \varepsilon_t$, with ε_t generated as either an NIID(0,1) sequence of variables or as an ARCH(1) process.

Specifically, results are reported in Tables S.1–S.10 for the following cases:[2]

- Tables S.1 and S.2, NIID: $\delta = 0$, $\phi = 0$, $\varepsilon_t \sim NIID(0, 1)$.

[1] We also considered the automatic bandwidth rule of Delgado and Robinson (1996) which does not assume that η_t follows an AR(1) process. We found that the tests based on this performed far worse than those based on m_{LR}, regardless of whether η_t was an AR(1) process or not. Consequently these results are not reported.

[2] We considered a much wider set of simulation DGPs for both size and power than are reported here, including higher-order ARFIMA processes; the broad conclusions drawn from these additional results were qualitatively similar to those drawn from the simulation DGPs reported here.

- Tables S.3 and S.4, AR(1): $\delta = 0$, $\phi = 0.5$, $\varepsilon_t \sim NIID(0,1)$.
- Tables S.5 and S.6, ARCH(1): $\delta = 0$, $\phi = 0$, $\varepsilon_t = \sigma_t z_t$, $z_t \sim NIID(0,1)$, $\sigma_t^2 = 1/(1-0.5) + 0.5\varepsilon_{t-1}^2$.
- Tables S.7 and S.8, ARFIMA(0,0.3,0): $\delta = 0.3$, $\phi = 0$, $\varepsilon_t \sim NIID(0,1)$.
- Tables S.9 and S.10, ARFIMA(0,−0.3,0): $\delta = -0.3$, $\phi = 0$, $\varepsilon_t \sim NIID(0,1)$.

In the case of the ARCH(1) model in Tables S.5 and S.6, the process was initiated at $\sigma_1^2 = 1$ such that the unconditional variance of ε_t is unity as in the NIID(0,1) case. The results reported in Tables S.1–S.6 relate to tests of the null hypothesis $H_0 : \delta = 0$, while the results in Tables S.7 and S.8 relate to tests of the null hypothesis $H_0 : \delta = 0.3$, and the results in Tables S.9 and S.10 relate to tests of the null hypothesis $H_0 : \delta = -0.3$. Notice that results for the KPSS-type tests are not reported in Tables S.7–S.10 as these are only appropriate for testing the null hypothesis $H_0 : \delta = 0$.

S.3 Simulation DGPs: Empirical Power

We investigate the empirical power properties of the tests by generating $u_t = \Delta^{-\delta}\varepsilon_t$ in (1) with ε_t being either an NIID(0,1) process or an ARCH(1) process. Results are reported in Tables S.11–S.18 for the following cases:

- Tables S.11 and S.12, NIID: $\delta = -0.15$, $\varepsilon_t \sim NIID(0,1)$.
- Tables S.13 and S.14, NIID: $\delta = 0.15$, $\varepsilon_t \sim NIID(0,1)$.
- Tables S.15 and S.16, ARCH(1): $\delta = -0.15$, $\varepsilon_t = \sigma_t z_t$, $z_t \sim NIID(0,1)$, $\sigma_t^2 = 1/(1-0.5) + 0.5\varepsilon_{t-1}^2$.
- Tables S.17 and S.18, ARCH(1): $\delta = 0.15$, $\varepsilon_t = \sigma_t z_t$, $z_t \sim NIID(0,1)$, $\sigma_t^2 = 1/(1-0.5) + 0.5\varepsilon_{t-1}^2$.

In each case we consider tests of the null hypothesis $H_0 : \delta = 0$. The KPSS-type tests are only reported in the DGPs with positive long memory, $\delta = 0.15$; in the DGPs with negative long memory, $\delta = -0.15$, no rejections are observed, as would be expected. For the same reason, only right-tailed t-tests are reported for $\delta = 0.15$ and only left-tailed t-tests for $\delta = -0.15$.

Moreover, results for the $LM_m(\delta_0)$, $t_m(\delta_0)$ and (where relevant) $KPSS$ tests that do not estimate a break, are reported only for $\beta_2 = 0$, as these tests are not size controlled (even asymptotically) when $\beta_2 \neq 0$.

Additional References

Delgado, M.A., and Robinson, P.M. (1996). Optimal spectral bandwidth for long memory. *Statistica Sinica* 6, 97–112.

Newey, W.K., and West, K.D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61, 631–653.

Table S.1: Empirical null rejection frequencies. DGP $u_t = \varepsilon_t$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0$. $T = 512$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m	$LM_m(\delta_0)$				$LM_m(\delta_0;\tau^*)$	$LM_m(\delta_0;\widehat{\tau})$			
$\lfloor T^{0.40} \rfloor = 12$	0.007	0.852	1.000	1.000	0.003	0.001	0.002	0.003	0.003
$\lfloor T^{0.45} \rfloor = 16$	0.010	0.908	1.000	1.000	0.004	0.003	0.003	0.005	0.005
$\lfloor T^{0.50} \rfloor = 22$	0.015	0.935	1.000	1.000	0.007	0.005	0.005	0.007	0.007
$\lfloor T^{0.55} \rfloor = 30$	0.016	0.953	1.000	1.000	0.008	0.009	0.009	0.009	0.008
$\lfloor T^{0.60} \rfloor = 42$	0.021	0.957	1.000	1.000	0.014	0.014	0.013	0.014	0.015
$\lfloor T^{0.65} \rfloor = 57$	0.027	0.956	1.000	1.000	0.021	0.022	0.019	0.020	0.021
$\lfloor T^{0.70} \rfloor = 78$	0.029	0.951	1.000	1.000	0.025	0.028	0.023	0.026	0.024
$\lfloor T^{0.75} \rfloor = 107$	0.034	0.942	1.000	1.000	0.029	0.032	0.029	0.029	0.029
$\lfloor T^{0.80} \rfloor = 147$	0.036	0.930	1.000	1.000	0.030	0.036	0.030	0.029	0.030
m_{LR}	0.029	0.916	1.000	1.000	0.025	0.029	0.027	0.026	0.026
m	$t_{m,L}(\delta_0)$				$t_{m,L}(\delta_0;\tau^*)$	$t_{m,L}(\delta_0;\widehat{\tau})$			
$\lfloor T^{0.40} \rfloor = 12$	0.002	0.000	0.000	0.000	0.003	0.006	0.003	0.003	0.003
$\lfloor T^{0.45} \rfloor = 16$	0.003	0.000	0.000	0.000	0.007	0.012	0.008	0.008	0.008
$\lfloor T^{0.50} \rfloor = 22$	0.006	0.000	0.000	0.000	0.013	0.024	0.014	0.012	0.012
$\lfloor T^{0.55} \rfloor = 30$	0.010	0.000	0.000	0.000	0.016	0.031	0.020	0.017	0.017
$\lfloor T^{0.60} \rfloor = 42$	0.014	0.000	0.000	0.000	0.023	0.043	0.029	0.025	0.023
$\lfloor T^{0.65} \rfloor = 57$	0.020	0.000	0.000	0.000	0.030	0.056	0.038	0.032	0.030
$\lfloor T^{0.70} \rfloor = 78$	0.025	0.000	0.000	0.000	0.034	0.061	0.043	0.035	0.034
$\lfloor T^{0.75} \rfloor = 107$	0.028	0.000	0.000	0.000	0.039	0.066	0.049	0.042	0.038
$\lfloor T^{0.80} \rfloor = 147$	0.034	0.000	0.000	0.000	0.044	0.074	0.057	0.047	0.044
m_{LR}	0.027	0.000	0.000	0.000	0.038	0.065	0.047	0.040	0.038
m	$t_{m,R}(\delta_0)$				$t_{m,R}(\delta_0;\tau^*)$	$t_{m,R}(\delta_0;\widehat{\tau})$			
$\lfloor T^{0.40} \rfloor = 12$	0.017	0.919	1.000	1.000	0.006	0.002	0.004	0.007	0.007
$\lfloor T^{0.45} \rfloor = 16$	0.025	0.950	1.000	1.000	0.010	0.003	0.007	0.010	0.010
$\lfloor T^{0.50} \rfloor = 22$	0.028	0.964	1.000	1.000	0.013	0.004	0.010	0.012	0.012
$\lfloor T^{0.55} \rfloor = 30$	0.030	0.973	1.000	1.000	0.013	0.005	0.012	0.013	0.014
$\lfloor T^{0.60} \rfloor = 42$	0.038	0.976	1.000	1.000	0.018	0.006	0.012	0.017	0.018
$\lfloor T^{0.65} \rfloor = 57$	0.039	0.973	1.000	1.000	0.021	0.008	0.014	0.021	0.021
$\lfloor T^{0.70} \rfloor = 78$	0.045	0.970	1.000	1.000	0.026	0.010	0.018	0.026	0.026
$\lfloor T^{0.75} \rfloor = 107$	0.047	0.964	1.000	1.000	0.030	0.013	0.020	0.029	0.030
$\lfloor T^{0.80} \rfloor = 147$	0.048	0.954	1.000	1.000	0.031	0.013	0.021	0.029	0.031
m_{LR}	0.042	0.945	1.000	1.000	0.028	0.011	0.018	0.024	0.027
M	$KPSS$				$KPSS(\tau^*)$	$KPSS(\widehat{\tau})$			
M_e	0.051	0.999	1.000	1.000	0.049	0.038	0.042	0.046	0.048
M_{LR}	0.048	0.998	1.000	1.000	0.047	0.045	0.042	0.045	0.046

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$; $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.

Table S.2: Empirical null rejection frequencies. DGP $u_t = \varepsilon_t$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0$. $T = 1024$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m		$LM_m(\delta_0)$			$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 16$	0.008	0.999	1.000	1.000	0.004	0.003	0.004	0.004	0.004
$\lfloor T^{0.45} \rfloor = 22$	0.015	1.000	1.000	1.000	0.009	0.006	0.008	0.008	0.009
$\lfloor T^{0.50} \rfloor = 32$	0.021	1.000	1.000	1.000	0.014	0.013	0.013	0.014	0.015
$\lfloor T^{0.55} \rfloor = 45$	0.023	1.000	1.000	1.000	0.016	0.017	0.015	0.016	0.016
$\lfloor T^{0.60} \rfloor = 63$	0.025	1.000	1.000	1.000	0.020	0.020	0.019	0.019	0.020
$\lfloor T^{0.65} \rfloor = 90$	0.031	1.000	1.000	1.000	0.024	0.027	0.024	0.023	0.024
$\lfloor T^{0.70} \rfloor = 127$	0.034	1.000	1.000	1.000	0.030	0.030	0.031	0.030	0.030
$\lfloor T^{0.75} \rfloor = 181$	0.035	1.000	1.000	1.000	0.032	0.035	0.033	0.033	0.032
$\lfloor T^{0.80} \rfloor = 256$	0.037	0.999	1.000	1.000	0.034	0.038	0.035	0.035	0.034
m_{LR}	0.035	0.998	1.000	1.000	0.033	0.037	0.034	0.033	0.033
m		$t_{m,L}(\delta_0)$			$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 16$	0.004	0.000	0.000	0.000	0.009	0.013	0.008	0.008	0.008
$\lfloor T^{0.45} \rfloor = 22$	0.007	0.000	0.000	0.000	0.013	0.022	0.014	0.013	0.013
$\lfloor T^{0.50} \rfloor = 32$	0.012	0.000	0.000	0.000	0.019	0.035	0.020	0.019	0.019
$\lfloor T^{0.55} \rfloor = 45$	0.016	0.000	0.000	0.000	0.025	0.044	0.028	0.026	0.026
$\lfloor T^{0.60} \rfloor = 63$	0.019	0.000	0.000	0.000	0.028	0.050	0.030	0.029	0.028
$\lfloor T^{0.65} \rfloor = 90$	0.026	0.000	0.000	0.000	0.038	0.060	0.043	0.039	0.038
$\lfloor T^{0.70} \rfloor = 127$	0.028	0.000	0.000	0.000	0.040	0.065	0.046	0.041	0.040
$\lfloor T^{0.75} \rfloor = 181$	0.031	0.000	0.000	0.000	0.041	0.068	0.050	0.043	0.041
$\lfloor T^{0.80} \rfloor = 256$	0.034	0.000	0.000	0.000	0.045	0.074	0.056	0.050	0.046
m_{LR}	0.033	0.000	0.000	0.000	0.045	0.070	0.053	0.047	0.045
m		$t_{m,R}(\delta_0)$			$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 16$	0.022	0.999	1.000	1.000	0.010	0.003	0.009	0.009	0.010
$\lfloor T^{0.45} \rfloor = 22$	0.029	1.000	1.000	1.000	0.013	0.005	0.013	0.014	0.014
$\lfloor T^{0.50} \rfloor = 32$	0.034	1.000	1.000	1.000	0.018	0.006	0.016	0.018	0.018
$\lfloor T^{0.55} \rfloor = 45$	0.036	1.000	1.000	1.000	0.021	0.006	0.018	0.021	0.020
$\lfloor T^{0.60} \rfloor = 63$	0.038	1.000	1.000	1.000	0.022	0.008	0.020	0.023	0.023
$\lfloor T^{0.65} \rfloor = 90$	0.043	1.000	1.000	1.000	0.027	0.011	0.023	0.027	0.027
$\lfloor T^{0.70} \rfloor = 127$	0.045	1.000	1.000	1.000	0.030	0.014	0.026	0.029	0.030
$\lfloor T^{0.75} \rfloor = 181$	0.044	1.000	1.000	1.000	0.030	0.014	0.026	0.029	0.030
$\lfloor T^{0.80} \rfloor = 256$	0.047	0.999	1.000	1.000	0.033	0.017	0.027	0.031	0.032
m_{LR}	0.041	0.999	1.000	1.000	0.029	0.015	0.024	0.028	0.029
M		$KPSS$			$KPSS(\tau^*)$		$KPSS(\widehat{\tau})$		
M_e	0.052	1.000	1.000	1.000	0.046	0.042	0.045	0.046	0.047
M_{LR}	0.053	1.000	1.000	1.000	0.045	0.045	0.044	0.043	0.045

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$; $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.

Table S.3: Empirical null rejection frequencies. DGP $u_t = \phi u_{t-1} + \varepsilon_t$, $\phi = 0.5$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0$. $T = 512$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m		$LM_m(\delta_0)$			$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.008	0.205	0.862	1.000	0.003	0.001	0.001	0.002	0.003
$\lfloor T^{0.45} \rfloor = 16$	0.012	0.274	0.921	1.000	0.005	0.003	0.002	0.003	0.005
$\lfloor T^{0.50} \rfloor = 22$	0.020	0.333	0.954	1.000	0.008	0.004	0.004	0.005	0.007
$\lfloor T^{0.55} \rfloor = 30$	0.030	0.405	0.976	1.000	0.014	0.007	0.008	0.012	0.014
$\lfloor T^{0.60} \rfloor = 42$	0.072	0.523	0.990	1.000	0.040	0.015	0.019	0.031	0.038
$\lfloor T^{0.65} \rfloor = 57$	0.191	0.693	0.997	1.000	0.128	0.064	0.078	0.104	0.122
$\lfloor T^{0.70} \rfloor = 78$	0.539	0.899	1.000	1.000	0.448	0.311	0.345	0.399	0.434
$\lfloor T^{0.75} \rfloor = 107$	0.939	0.994	1.000	1.000	0.910	0.846	0.864	0.890	0.902
$\lfloor T^{0.80} \rfloor = 147$	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000
m_{LR}	0.042	0.432	0.976	1.000	0.022	0.010	0.011	0.017	0.021
m		$t_{m,L}(\delta_0)$			$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.001	0.000	0.000	0.000	0.002	0.006	0.004	0.003	0.003
$\lfloor T^{0.45} \rfloor = 16$	0.002	0.000	0.000	0.000	0.006	0.010	0.009	0.006	0.006
$\lfloor T^{0.50} \rfloor = 22$	0.004	0.000	0.000	0.000	0.008	0.015	0.012	0.009	0.008
$\lfloor T^{0.55} \rfloor = 30$	0.003	0.000	0.000	0.000	0.006	0.013	0.011	0.007	0.006
$\lfloor T^{0.60} \rfloor = 42$	0.002	0.000	0.000	0.000	0.003	0.007	0.005	0.004	0.003
$\lfloor T^{0.65} \rfloor = 57$	0.001	0.000	0.000	0.000	0.001	0.002	0.001	0.001	0.001
$\lfloor T^{0.70} \rfloor = 78$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\lfloor T^{0.75} \rfloor = 107$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\lfloor T^{0.80} \rfloor = 147$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
m_{LR}	0.002	0.000	0.000	0.000	0.005	0.010	0.008	0.006	0.005
m		$t_{m,R}(\delta_0)$			$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.019	0.310	0.926	1.000	0.007	0.002	0.003	0.005	0.007
$\lfloor T^{0.45} \rfloor = 16$	0.029	0.376	0.959	1.000	0.012	0.003	0.005	0.008	0.011
$\lfloor T^{0.50} \rfloor = 22$	0.040	0.430	0.976	1.000	0.017	0.005	0.008	0.013	0.016
$\lfloor T^{0.55} \rfloor = 30$	0.063	0.502	0.987	1.000	0.031	0.010	0.015	0.025	0.029
$\lfloor T^{0.60} \rfloor = 42$	0.123	0.621	0.993	1.000	0.075	0.033	0.041	0.059	0.071
$\lfloor T^{0.65} \rfloor = 57$	0.291	0.775	0.999	1.000	0.209	0.114	0.137	0.173	0.200
$\lfloor T^{0.70} \rfloor = 78$	0.653	0.939	1.000	1.000	0.572	0.436	0.468	0.526	0.558
$\lfloor T^{0.75} \rfloor = 107$	0.969	0.997	1.000	1.000	0.952	0.913	0.924	0.942	0.947
$\lfloor T^{0.80} \rfloor = 147$	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000
m_{LR}	0.081	0.531	0.987	1.000	0.045	0.017	0.024	0.038	0.045
M		$KPSS$			$KPSS(\tau^*)$		$KPSS(\widehat{\tau})$		
M_e	0.111	0.820	1.000	1.000	0.124	0.092	0.083	0.103	0.119
M_{LR}	0.064	0.728	0.999	1.000	0.062	0.052	0.048	0.054	0.058

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$; $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.

Table S.4: Empirical null rejection frequencies. DGP $u_t = \phi u_{t-1} + \varepsilon_t$, $\phi = 0.5$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0$. $T = 1024$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m		$LM_m(\delta_0)$			$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.008	0.205	0.862	1.000	0.003	0.001	0.001	0.002	0.003
$\lfloor T^{0.45} \rfloor = 16$	0.012	0.274	0.921	1.000	0.005	0.003	0.002	0.003	0.005
$\lfloor T^{0.50} \rfloor = 22$	0.020	0.333	0.954	1.000	0.008	0.004	0.004	0.005	0.007
$\lfloor T^{0.55} \rfloor = 30$	0.030	0.405	0.976	1.000	0.014	0.007	0.008	0.012	0.014
$\lfloor T^{0.60} \rfloor = 42$	0.072	0.523	0.990	1.000	0.040	0.015	0.019	0.031	0.038
$\lfloor T^{0.65} \rfloor = 57$	0.191	0.693	0.997	1.000	0.128	0.064	0.078	0.104	0.122
$\lfloor T^{0.70} \rfloor = 78$	0.539	0.899	1.000	1.000	0.448	0.311	0.345	0.399	0.434
$\lfloor T^{0.75} \rfloor = 107$	0.939	0.994	1.000	1.000	0.910	0.846	0.864	0.890	0.902
$\lfloor T^{0.80} \rfloor = 147$	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000
m_{LR}	0.042	0.432	0.976	1.000	0.022	0.010	0.011	0.017	0.021
m		$t_{m,L}(\delta_0)$			$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.001	0.000	0.000	0.000	0.002	0.006	0.004	0.003	0.003
$\lfloor T^{0.45} \rfloor = 16$	0.002	0.000	0.000	0.000	0.006	0.010	0.009	0.006	0.006
$\lfloor T^{0.50} \rfloor = 22$	0.004	0.000	0.000	0.000	0.008	0.015	0.012	0.009	0.008
$\lfloor T^{0.55} \rfloor = 30$	0.003	0.000	0.000	0.000	0.006	0.013	0.011	0.007	0.006
$\lfloor T^{0.60} \rfloor = 42$	0.002	0.000	0.000	0.000	0.003	0.007	0.005	0.004	0.003
$\lfloor T^{0.65} \rfloor = 57$	0.001	0.000	0.000	0.000	0.001	0.002	0.001	0.001	0.001
$\lfloor T^{0.70} \rfloor = 78$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\lfloor T^{0.75} \rfloor = 107$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\lfloor T^{0.80} \rfloor = 147$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
m_{LR}	0.002	0.000	0.000	0.000	0.005	0.010	0.008	0.006	0.005
m		$t_{m,R}(\delta_0)$			$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.019	0.310	0.926	1.000	0.007	0.002	0.003	0.005	0.007
$\lfloor T^{0.45} \rfloor = 16$	0.029	0.376	0.959	1.000	0.012	0.003	0.005	0.008	0.011
$\lfloor T^{0.50} \rfloor = 22$	0.040	0.430	0.976	1.000	0.017	0.005	0.008	0.013	0.016
$\lfloor T^{0.55} \rfloor = 30$	0.063	0.502	0.987	1.000	0.031	0.010	0.015	0.025	0.029
$\lfloor T^{0.60} \rfloor = 42$	0.123	0.621	0.993	1.000	0.075	0.033	0.041	0.059	0.071
$\lfloor T^{0.65} \rfloor = 57$	0.291	0.775	0.999	1.000	0.209	0.114	0.137	0.173	0.200
$\lfloor T^{0.70} \rfloor = 78$	0.653	0.939	1.000	1.000	0.572	0.436	0.468	0.526	0.558
$\lfloor T^{0.75} \rfloor = 107$	0.969	0.997	1.000	1.000	0.952	0.913	0.924	0.942	0.947
$\lfloor T^{0.80} \rfloor = 147$	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000
m_{LR}	0.081	0.531	0.987	1.000	0.045	0.017	0.024	0.038	0.045
M		$KPSS$			$KPSS(\tau^*)$		$KPSS(\widehat{\tau})$		
M_e	0.111	0.820	1.000	1.000	0.124	0.092	0.083	0.103	0.119
M_{LR}	0.064	0.728	0.999	1.000	0.062	0.052	0.048	0.054	0.058

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$; $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.

Table S.5: Empirical null rejection frequencies. DGP $u_t = \varepsilon_t = \sigma_t z_t$, $z_t \sim NIID(0,1)$, $\sigma_t^2 = 1/(1-0.5) + 0.5\varepsilon_{t-1}^2$. Tests of $H_0 : \delta = 0$. $T = 512$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m		$LM_m(\delta_0)$			$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.007	0.848	0.999	1.000	0.003	0.000	0.002	0.003	0.003
$\lfloor T^{0.45} \rfloor = 16$	0.010	0.899	1.000	1.000	0.004	0.002	0.003	0.004	0.004
$\lfloor T^{0.50} \rfloor = 22$	0.015	0.931	1.000	1.000	0.008	0.004	0.006	0.008	0.008
$\lfloor T^{0.55} \rfloor = 30$	0.016	0.945	1.000	1.000	0.010	0.008	0.009	0.010	0.011
$\lfloor T^{0.60} \rfloor = 42$	0.022	0.948	1.000	1.000	0.014	0.016	0.012	0.014	0.014
$\lfloor T^{0.65} \rfloor = 57$	0.033	0.945	1.000	1.000	0.025	0.026	0.023	0.024	0.027
$\lfloor T^{0.70} \rfloor = 78$	0.043	0.939	1.000	1.000	0.038	0.041	0.035	0.037	0.037
$\lfloor T^{0.75} \rfloor = 107$	0.065	0.930	1.000	1.000	0.058	0.065	0.057	0.058	0.058
$\lfloor T^{0.80} \rfloor = 147$	0.096	0.910	1.000	1.000	0.092	0.101	0.091	0.091	0.090
m_{LR}	0.071	0.895	1.000	1.000	0.069	0.075	0.068	0.067	0.068
m		$t_{m,L}(\delta_0)$			$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.001	0.000	0.000	0.000	0.003	0.005	0.003	0.003	0.003
$\lfloor T^{0.45} \rfloor = 16$	0.002	0.000	0.000	0.000	0.005	0.011	0.007	0.006	0.006
$\lfloor T^{0.50} \rfloor = 22$	0.008	0.000	0.000	0.000	0.013	0.022	0.014	0.013	0.012
$\lfloor T^{0.55} \rfloor = 30$	0.011	0.000	0.000	0.000	0.018	0.029	0.020	0.018	0.018
$\lfloor T^{0.60} \rfloor = 42$	0.016	0.000	0.000	0.000	0.025	0.041	0.030	0.026	0.026
$\lfloor T^{0.65} \rfloor = 57$	0.026	0.000	0.000	0.000	0.037	0.057	0.043	0.038	0.037
$\lfloor T^{0.70} \rfloor = 78$	0.037	0.000	0.000	0.000	0.049	0.075	0.055	0.051	0.049
$\lfloor T^{0.75} \rfloor = 107$	0.054	0.000	0.000	0.000	0.069	0.098	0.079	0.072	0.067
$\lfloor T^{0.80} \rfloor = 147$	0.078	0.001	0.000	0.000	0.096	0.132	0.108	0.100	0.095
m_{LR}	0.068	0.001	0.000	0.000	0.083	0.113	0.094	0.087	0.083
m		$t_{m,R}(\delta_0)$			$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.016	0.911	1.000	1.000	0.007	0.002	0.005	0.006	0.007
$\lfloor T^{0.45} \rfloor = 16$	0.023	0.942	1.000	1.000	0.010	0.003	0.008	0.009	0.010
$\lfloor T^{0.50} \rfloor = 22$	0.028	0.958	1.000	1.000	0.012	0.004	0.009	0.012	0.013
$\lfloor T^{0.55} \rfloor = 30$	0.030	0.962	1.000	1.000	0.015	0.005	0.012	0.015	0.016
$\lfloor T^{0.60} \rfloor = 42$	0.037	0.967	1.000	1.000	0.020	0.007	0.015	0.019	0.019
$\lfloor T^{0.65} \rfloor = 57$	0.044	0.964	1.000	1.000	0.025	0.011	0.018	0.023	0.025
$\lfloor T^{0.70} \rfloor = 78$	0.052	0.961	1.000	1.000	0.033	0.016	0.025	0.031	0.032
$\lfloor T^{0.75} \rfloor = 107$	0.065	0.952	1.000	1.000	0.047	0.025	0.038	0.044	0.047
$\lfloor T^{0.80} \rfloor = 147$	0.085	0.935	1.000	1.000	0.065	0.042	0.054	0.061	0.064
m_{LR}	0.061	0.923	1.000	1.000	0.044	0.025	0.034	0.042	0.044
M		$KPSS$			$KPSS(\tau^*)$		$KPSS(\widehat{\tau})$		
M_e	0.051	0.996	1.000	1.000	0.050	0.042	0.041	0.046	0.049
M_{LR}	0.047	0.995	1.000	1.000	0.047	0.050	0.040	0.043	0.046

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$; $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.

Table S.6: Empirical null rejection frequencies. DGP $u_t = \varepsilon_t = \sigma_t z_t$, $z_t \sim NIID(0,1)$, $\sigma_t^2 = 1/(1-0.5) + 0.5\varepsilon_{t-1}^2$. Tests of $H_0 : \delta = 0$. $T = 1024$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m		$LM_m(\delta_0)$			$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 16$	0.009	0.997	1.000	1.000	0.004	0.003	0.004	0.004	0.004
$\lfloor T^{0.45} \rfloor = 22$	0.014	0.998	1.000	1.000	0.008	0.005	0.007	0.009	0.008
$\lfloor T^{0.50} \rfloor = 32$	0.021	0.999	1.000	1.000	0.013	0.010	0.012	0.013	0.013
$\lfloor T^{0.55} \rfloor = 45$	0.022	0.999	1.000	1.000	0.016	0.015	0.016	0.015	0.015
$\lfloor T^{0.60} \rfloor = 63$	0.024	0.999	1.000	1.000	0.020	0.021	0.019	0.019	0.019
$\lfloor T^{0.65} \rfloor = 90$	0.033	0.999	1.000	1.000	0.029	0.030	0.028	0.028	0.029
$\lfloor T^{0.70} \rfloor = 127$	0.044	0.999	1.000	1.000	0.041	0.044	0.042	0.042	0.042
$\lfloor T^{0.75} \rfloor = 181$	0.065	0.998	1.000	1.000	0.062	0.068	0.063	0.061	0.063
$\lfloor T^{0.80} \rfloor = 256$	0.097	0.994	1.000	1.000	0.093	0.100	0.094	0.094	0.094
m_{LR}	0.086	0.993	1.000	1.000	0.084	0.092	0.085	0.086	0.085
m		$t_{m,L}(\delta_0)$			$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 16$	0.004	0.000	0.000	0.000	0.007	0.012	0.008	0.007	0.007
$\lfloor T^{0.45} \rfloor = 22$	0.005	0.000	0.000	0.000	0.009	0.019	0.010	0.010	0.009
$\lfloor T^{0.50} \rfloor = 32$	0.009	0.000	0.000	0.000	0.018	0.034	0.019	0.018	0.018
$\lfloor T^{0.55} \rfloor = 45$	0.014	0.000	0.000	0.000	0.023	0.042	0.025	0.023	0.023
$\lfloor T^{0.60} \rfloor = 63$	0.018	0.000	0.000	0.000	0.029	0.050	0.032	0.030	0.029
$\lfloor T^{0.65} \rfloor = 90$	0.029	0.000	0.000	0.000	0.038	0.060	0.041	0.040	0.037
$\lfloor T^{0.70} \rfloor = 127$	0.039	0.000	0.000	0.000	0.052	0.077	0.058	0.054	0.052
$\lfloor T^{0.75} \rfloor = 181$	0.056	0.000	0.000	0.000	0.069	0.096	0.074	0.070	0.070
$\lfloor T^{0.80} \rfloor = 256$	0.076	0.000	0.000	0.000	0.094	0.124	0.102	0.095	0.095
m_{LR}	0.078	0.000	0.000	0.000	0.093	0.125	0.103	0.096	0.093
m		$t_{m,R}(\delta_0)$			$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 16$	0.019	0.998	1.000	1.000	0.008	0.002	0.007	0.009	0.008
$\lfloor T^{0.45} \rfloor = 22$	0.029	0.999	1.000	1.000	0.015	0.005	0.012	0.014	0.015
$\lfloor T^{0.50} \rfloor = 32$	0.033	0.999	1.000	1.000	0.018	0.006	0.015	0.017	0.019
$\lfloor T^{0.55} \rfloor = 45$	0.036	0.999	1.000	1.000	0.019	0.007	0.016	0.019	0.019
$\lfloor T^{0.60} \rfloor = 63$	0.038	0.999	1.000	1.000	0.023	0.009	0.019	0.022	0.023
$\lfloor T^{0.65} \rfloor = 90$	0.043	0.999	1.000	1.000	0.026	0.011	0.024	0.026	0.026
$\lfloor T^{0.70} \rfloor = 127$	0.053	0.999	1.000	1.000	0.035	0.018	0.030	0.035	0.034
$\lfloor T^{0.75} \rfloor = 181$	0.063	0.999	1.000	1.000	0.047	0.028	0.041	0.046	0.048
$\lfloor T^{0.80} \rfloor = 256$	0.083	0.996	1.000	1.000	0.067	0.045	0.060	0.065	0.067
m_{LR}	0.069	0.995	1.000	1.000	0.053	0.035	0.047	0.053	0.054
M		$KPSS$			$KPSS(\tau^*)$		$KPSS(\widehat{\tau})$		
M_e	0.054	1.000	1.000	1.000	0.048	0.043	0.043	0.046	0.048
M_{LR}	0.051	1.000	1.000	1.000	0.045	0.047	0.042	0.043	0.045

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$; $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.

Table S.7: Empirical null rejection frequencies. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = 0.3$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0.3$. $T = 512$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m		$LM_m(\delta_0)$			$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.008	0.043	0.210	0.807	0.002	0.001	0.001	0.001	0.002
$\lfloor T^{0.45} \rfloor = 16$	0.011	0.061	0.280	0.902	0.003	0.003	0.002	0.001	0.003
$\lfloor T^{0.50} \rfloor = 22$	0.018	0.073	0.334	0.949	0.006	0.005	0.005	0.004	0.004
$\lfloor T^{0.55} \rfloor = 30$	0.020	0.085	0.366	0.968	0.009	0.011	0.009	0.007	0.007
$\lfloor T^{0.60} \rfloor = 42$	0.024	0.087	0.381	0.977	0.013	0.017	0.014	0.012	0.012
$\lfloor T^{0.65} \rfloor = 57$	0.029	0.093	0.393	0.980	0.020	0.028	0.023	0.019	0.019
$\lfloor T^{0.70} \rfloor = 78$	0.032	0.091	0.389	0.979	0.025	0.035	0.031	0.025	0.024
$\lfloor T^{0.75} \rfloor = 107$	0.035	0.086	0.358	0.979	0.031	0.043	0.040	0.034	0.030
$\lfloor T^{0.80} \rfloor = 147$	0.040	0.071	0.303	0.969	0.040	0.066	0.061	0.049	0.042
m_{LR}	0.029	0.058	0.278	0.963	0.029	0.051	0.046	0.037	0.032
m		$t_{m,L}(\delta_0)$			$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.001	0.000	0.000	0.000	0.003	0.009	0.007	0.005	0.004
$\lfloor T^{0.45} \rfloor = 16$	0.003	0.002	0.000	0.000	0.007	0.018	0.016	0.013	0.009
$\lfloor T^{0.50} \rfloor = 22$	0.005	0.002	0.000	0.000	0.012	0.031	0.028	0.019	0.014
$\lfloor T^{0.55} \rfloor = 30$	0.009	0.003	0.000	0.000	0.017	0.042	0.037	0.026	0.020
$\lfloor T^{0.60} \rfloor = 42$	0.012	0.005	0.001	0.000	0.023	0.056	0.049	0.036	0.027
$\lfloor T^{0.65} \rfloor = 57$	0.022	0.009	0.001	0.000	0.034	0.073	0.064	0.049	0.037
$\lfloor T^{0.70} \rfloor = 78$	0.024	0.013	0.002	0.000	0.039	0.082	0.072	0.056	0.043
$\lfloor T^{0.75} \rfloor = 107$	0.033	0.017	0.002	0.000	0.050	0.094	0.085	0.069	0.055
$\lfloor T^{0.80} \rfloor = 147$	0.053	0.029	0.004	0.000	0.072	0.128	0.121	0.101	0.078
m_{LR}	0.045	0.024	0.004	0.000	0.062	0.114	0.104	0.087	0.070
m		$t_{m,R}(\delta_0)$			$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.019	0.080	0.314	0.898	0.005	0.001	0.001	0.001	0.003
$\lfloor T^{0.45} \rfloor = 16$	0.027	0.103	0.379	0.950	0.007	0.001	0.002	0.002	0.005
$\lfloor T^{0.50} \rfloor = 22$	0.032	0.118	0.435	0.973	0.010	0.001	0.002	0.004	0.007
$\lfloor T^{0.55} \rfloor = 30$	0.037	0.126	0.464	0.984	0.013	0.002	0.003	0.006	0.010
$\lfloor T^{0.60} \rfloor = 42$	0.041	0.133	0.480	0.986	0.017	0.003	0.005	0.007	0.013
$\lfloor T^{0.65} \rfloor = 57$	0.042	0.137	0.482	0.989	0.020	0.005	0.007	0.009	0.015
$\lfloor T^{0.70} \rfloor = 78$	0.047	0.131	0.476	0.988	0.023	0.006	0.008	0.011	0.019
$\lfloor T^{0.75} \rfloor = 107$	0.044	0.123	0.445	0.988	0.023	0.007	0.009	0.013	0.018
$\lfloor T^{0.80} \rfloor = 147$	0.036	0.097	0.388	0.981	0.018	0.005	0.007	0.009	0.014
m_{LR}	0.029	0.083	0.360	0.978	0.015	0.005	0.005	0.007	0.012

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0.3$; $t_{m,R}$ denotes alternative $H_1 : \delta > 0.3$.

Table S.8: Empirical null rejection frequencies. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = 0.3$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0.3$. $T = 1024$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m	$LM_m(\delta_0)$				$LM_m(\delta_0;\tau^*)$	$LM_m(\delta_0;\widehat{\tau})$			
$\lfloor T^{0.40} \rfloor = 16$	0.010	0.080	0.374	0.968	0.004	0.004	0.003	0.003	0.003
$\lfloor T^{0.45} \rfloor = 22$	0.017	0.097	0.439	0.987	0.007	0.007	0.005	0.005	0.006
$\lfloor T^{0.50} \rfloor = 32$	0.024	0.109	0.486	0.994	0.013	0.015	0.012	0.012	0.011
$\lfloor T^{0.55} \rfloor = 45$	0.027	0.114	0.514	0.996	0.016	0.020	0.019	0.014	0.014
$\lfloor T^{0.60} \rfloor = 63$	0.029	0.116	0.525	0.997	0.019	0.024	0.021	0.018	0.018
$\lfloor T^{0.65} \rfloor = 90$	0.033	0.117	0.513	0.997	0.024	0.033	0.029	0.023	0.023
$\lfloor T^{0.70} \rfloor = 127$	0.037	0.110	0.489	0.998	0.030	0.038	0.035	0.032	0.031
$\lfloor T^{0.75} \rfloor = 181$	0.036	0.103	0.445	0.997	0.032	0.046	0.041	0.035	0.033
$\lfloor T^{0.80} \rfloor = 256$	0.039	0.082	0.379	0.993	0.041	0.067	0.060	0.051	0.044
m_{LR}	0.038	0.072	0.338	0.989	0.041	0.066	0.062	0.054	0.042
m	$t_{m,L}(\delta_0)$				$t_{m,L}(\delta_0;\tau^*)$	$t_{m,L}(\delta_0;\widehat{\tau})$			
$\lfloor T^{0.40} \rfloor = 16$	0.004	0.001	0.000	0.000	0.008	0.018	0.015	0.011	0.009
$\lfloor T^{0.45} \rfloor = 22$	0.007	0.002	0.000	0.000	0.013	0.032	0.025	0.019	0.014
$\lfloor T^{0.50} \rfloor = 32$	0.012	0.004	0.000	0.000	0.020	0.045	0.038	0.027	0.021
$\lfloor T^{0.55} \rfloor = 45$	0.015	0.005	0.000	0.000	0.025	0.056	0.047	0.035	0.028
$\lfloor T^{0.60} \rfloor = 63$	0.017	0.006	0.000	0.000	0.029	0.061	0.053	0.037	0.031
$\lfloor T^{0.65} \rfloor = 90$	0.025	0.010	0.001	0.000	0.039	0.074	0.063	0.050	0.041
$\lfloor T^{0.70} \rfloor = 127$	0.029	0.012	0.001	0.000	0.042	0.083	0.072	0.058	0.045
$\lfloor T^{0.75} \rfloor = 181$	0.037	0.015	0.002	0.000	0.051	0.097	0.084	0.068	0.055
$\lfloor T^{0.80} \rfloor = 256$	0.053	0.026	0.004	0.000	0.073	0.128	0.118	0.097	0.077
m_{LR}	0.057	0.030	0.004	0.000	0.077	0.131	0.119	0.099	0.082
m	$t_{m,R}(\delta_0)$				$t_{m,R}(\delta_0;\tau^*)$	$t_{m,R}(\delta_0;\widehat{\tau})$			
$\lfloor T^{0.40} \rfloor = 16$	0.024	0.128	0.490	0.986	0.006	0.001	0.002	0.003	0.005
$\lfloor T^{0.45} \rfloor = 22$	0.033	0.145	0.546	0.995	0.012	0.003	0.005	0.006	0.010
$\lfloor T^{0.50} \rfloor = 32$	0.040	0.158	0.584	0.997	0.016	0.004	0.007	0.010	0.014
$\lfloor T^{0.55} \rfloor = 45$	0.043	0.165	0.602	0.998	0.019	0.004	0.006	0.011	0.016
$\lfloor T^{0.60} \rfloor = 63$	0.045	0.171	0.613	0.998	0.021	0.006	0.007	0.012	0.018
$\lfloor T^{0.65} \rfloor = 90$	0.049	0.168	0.602	0.999	0.026	0.006	0.009	0.015	0.022
$\lfloor T^{0.70} \rfloor = 127$	0.049	0.160	0.575	0.999	0.028	0.008	0.011	0.018	0.026
$\lfloor T^{0.75} \rfloor = 181$	0.043	0.143	0.532	0.999	0.026	0.008	0.011	0.016	0.023
$\lfloor T^{0.80} \rfloor = 256$	0.033	0.111	0.466	0.996	0.020	0.007	0.008	0.013	0.018
m_{LR}	0.030	0.095	0.424	0.994	0.018	0.006	0.007	0.011	0.016

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0.3$; $t_{m,R}$ denotes alternative $H_1 : \delta > 0.3$.

Table S.9: Empirical null rejection frequencies. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = -0.3$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = -0.3$. $T = 512$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m		$LM_m(\delta_0)$			$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.011	1.000	1.000	1.000	0.007	0.007	0.014	0.011	0.009
$\lfloor T^{0.45} \rfloor = 16$	0.017	1.000	1.000	1.000	0.012	0.011	0.019	0.018	0.015
$\lfloor T^{0.50} \rfloor = 22$	0.021	1.000	1.000	1.000	0.015	0.015	0.025	0.020	0.019
$\lfloor T^{0.55} \rfloor = 30$	0.025	1.000	1.000	1.000	0.017	0.020	0.031	0.026	0.022
$\lfloor T^{0.60} \rfloor = 42$	0.030	1.000	1.000	1.000	0.022	0.023	0.035	0.030	0.028
$\lfloor T^{0.65} \rfloor = 57$	0.035	1.000	1.000	1.000	0.027	0.026	0.042	0.038	0.036
$\lfloor T^{0.70} \rfloor = 78$	0.041	1.000	1.000	1.000	0.035	0.033	0.046	0.047	0.042
$\lfloor T^{0.75} \rfloor = 107$	0.050	1.000	1.000	1.000	0.042	0.041	0.053	0.053	0.050
$\lfloor T^{0.80} \rfloor = 147$	0.057	1.000	1.000	1.000	0.050	0.047	0.063	0.065	0.061
m_{LR}	0.055	1.000	1.000	1.000	0.044	0.043	0.056	0.058	0.055
m		$t_{m,L}(\delta_0)$			$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.001	0.000	0.000	0.000	0.002	0.001	0.001	0.001	0.001
$\lfloor T^{0.45} \rfloor = 16$	0.003	0.000	0.000	0.000	0.005	0.002	0.004	0.004	0.004
$\lfloor T^{0.50} \rfloor = 22$	0.005	0.000	0.000	0.000	0.007	0.005	0.006	0.006	0.006
$\lfloor T^{0.55} \rfloor = 30$	0.008	0.000	0.000	0.000	0.012	0.009	0.009	0.009	0.010
$\lfloor T^{0.60} \rfloor = 42$	0.011	0.000	0.000	0.000	0.016	0.013	0.013	0.013	0.015
$\lfloor T^{0.65} \rfloor = 57$	0.017	0.000	0.000	0.000	0.022	0.019	0.018	0.018	0.019
$\lfloor T^{0.70} \rfloor = 78$	0.019	0.000	0.000	0.000	0.022	0.023	0.019	0.018	0.019
$\lfloor T^{0.75} \rfloor = 107$	0.019	0.000	0.000	0.000	0.024	0.023	0.022	0.021	0.021
$\lfloor T^{0.80} \rfloor = 147$	0.015	0.000	0.000	0.000	0.019	0.021	0.017	0.017	0.017
m_{LR}	0.012	0.000	0.000	0.000	0.015	0.016	0.013	0.013	0.014
m		$t_{m,R}(\delta_0)$			$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.024	1.000	1.000	1.000	0.016	0.020	0.025	0.022	0.019
$\lfloor T^{0.45} \rfloor = 16$	0.034	1.000	1.000	1.000	0.024	0.026	0.036	0.033	0.028
$\lfloor T^{0.50} \rfloor = 22$	0.040	1.000	1.000	1.000	0.025	0.032	0.042	0.038	0.032
$\lfloor T^{0.55} \rfloor = 30$	0.044	1.000	1.000	1.000	0.032	0.036	0.052	0.046	0.039
$\lfloor T^{0.60} \rfloor = 42$	0.055	1.000	1.000	1.000	0.039	0.039	0.058	0.054	0.048
$\lfloor T^{0.65} \rfloor = 57$	0.055	1.000	1.000	1.000	0.042	0.043	0.061	0.058	0.054
$\lfloor T^{0.70} \rfloor = 78$	0.062	1.000	1.000	1.000	0.048	0.049	0.066	0.068	0.060
$\lfloor T^{0.75} \rfloor = 107$	0.073	1.000	1.000	1.000	0.059	0.058	0.076	0.077	0.073
$\lfloor T^{0.80} \rfloor = 147$	0.093	1.000	1.000	1.000	0.077	0.073	0.094	0.101	0.093
m_{LR}	0.084	1.000	1.000	1.000	0.073	0.069	0.088	0.092	0.088

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < -0.3$; $t_{m,R}$ denotes alternative $H_1 : \delta > -0.3$.

Table S.10: Empirical null rejection frequencies. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = -0.3$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = -0.3$. $T = 1024$. Nominal asymptotic 5% level tests.

β_2	0	0.5	1	2		0	0.5	1	2
m		$LM_m(\delta_0)$			$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 16$	0.014	1.000	1.000	1.000	0.011	0.016	0.016	0.014	0.013
$\lfloor T^{0.45} \rfloor = 22$	0.022	1.000	1.000	1.000	0.017	0.021	0.024	0.023	0.019
$\lfloor T^{0.50} \rfloor = 32$	0.028	1.000	1.000	1.000	0.023	0.025	0.033	0.030	0.027
$\lfloor T^{0.55} \rfloor = 45$	0.031	1.000	1.000	1.000	0.026	0.028	0.039	0.034	0.030
$\lfloor T^{0.60} \rfloor = 63$	0.034	1.000	1.000	1.000	0.029	0.030	0.041	0.039	0.033
$\lfloor T^{0.65} \rfloor = 90$	0.040	1.000	1.000	1.000	0.034	0.034	0.047	0.045	0.039
$\lfloor T^{0.70} \rfloor = 127$	0.044	1.000	1.000	1.000	0.041	0.041	0.054	0.053	0.047
$\lfloor T^{0.75} \rfloor = 181$	0.047	1.000	1.000	1.000	0.044	0.042	0.056	0.056	0.051
$\lfloor T^{0.80} \rfloor = 256$	0.059	1.000	1.000	1.000	0.052	0.050	0.065	0.064	0.058
m_{LR}	0.055	1.000	1.000	1.000	0.046	0.049	0.058	0.059	0.055
m		$t_{m,L}(\delta_0)$			$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\tau^*)$		
$\lfloor T^{0.40} \rfloor = 16$	0.004	0.000	0.000	0.000	0.006	0.003	0.004	0.005	0.005
$\lfloor T^{0.45} \rfloor = 22$	0.006	0.000	0.000	0.000	0.008	0.004	0.006	0.007	0.007
$\lfloor T^{0.50} \rfloor = 32$	0.010	0.000	0.000	0.000	0.013	0.009	0.011	0.011	0.012
$\lfloor T^{0.55} \rfloor = 45$	0.015	0.000	0.000	0.000	0.018	0.013	0.016	0.016	0.017
$\lfloor T^{0.60} \rfloor = 63$	0.016	0.000	0.000	0.000	0.021	0.017	0.017	0.017	0.019
$\lfloor T^{0.65} \rfloor = 90$	0.021	0.000	0.000	0.000	0.026	0.024	0.022	0.022	0.023
$\lfloor T^{0.70} \rfloor = 127$	0.023	0.000	0.000	0.000	0.028	0.024	0.024	0.024	0.025
$\lfloor T^{0.75} \rfloor = 181$	0.022	0.000	0.000	0.000	0.026	0.024	0.024	0.023	0.024
$\lfloor T^{0.80} \rfloor = 256$	0.016	0.000	0.000	0.000	0.020	0.020	0.019	0.017	0.018
m_{LR}	0.014	0.000	0.000	0.000	0.017	0.017	0.016	0.016	0.016
m		$t_{m,R}(\delta_0)$			$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 16$	0.031	1.000	1.000	1.000	0.022	0.032	0.029	0.028	0.025
$\lfloor T^{0.45} \rfloor = 22$	0.040	1.000	1.000	1.000	0.031	0.040	0.043	0.038	0.033
$\lfloor T^{0.50} \rfloor = 32$	0.048	1.000	1.000	1.000	0.039	0.044	0.052	0.046	0.044
$\lfloor T^{0.55} \rfloor = 45$	0.053	1.000	1.000	1.000	0.041	0.047	0.059	0.054	0.049
$\lfloor T^{0.60} \rfloor = 63$	0.051	1.000	1.000	1.000	0.041	0.046	0.061	0.056	0.048
$\lfloor T^{0.65} \rfloor = 90$	0.060	1.000	1.000	1.000	0.049	0.054	0.066	0.066	0.058
$\lfloor T^{0.70} \rfloor = 127$	0.064	1.000	1.000	1.000	0.053	0.057	0.073	0.069	0.063
$\lfloor T^{0.75} \rfloor = 181$	0.072	1.000	1.000	1.000	0.061	0.063	0.080	0.079	0.072
$\lfloor T^{0.80} \rfloor = 256$	0.090	1.000	1.000	1.000	0.078	0.082	0.096	0.096	0.089
m_{LR}	0.090	1.000	1.000	1.000	0.078	0.079	0.097	0.096	0.088

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < -0.3$; $t_{m,R}$ denotes alternative $H_1 : \delta > -0.3$.

Table S.11: Empirical power. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = -0.15$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0$. $T = 512$. Nominal asymptotic 5% level tests.

β_2	0		0	0.5	1	2
m	$LM_m(\delta_0)$	$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.001	0.001	0.001	0.002	0.001	0.001
$\lfloor T^{0.45} \rfloor = 16$	0.004	0.006	0.007	0.005	0.006	0.005
$\lfloor T^{0.50} \rfloor = 22$	0.013	0.020	0.030	0.019	0.019	0.019
$\lfloor T^{0.55} \rfloor = 30$	0.040	0.054	0.073	0.052	0.051	0.051
$\lfloor T^{0.60} \rfloor = 42$	0.107	0.129	0.165	0.126	0.125	0.125
$\lfloor T^{0.65} \rfloor = 57$	0.220	0.255	0.312	0.259	0.247	0.247
$\lfloor T^{0.70} \rfloor = 78$	0.392	0.426	0.489	0.436	0.419	0.418
$\lfloor T^{0.75} \rfloor = 107$	0.597	0.626	0.685	0.639	0.621	0.618
$\lfloor T^{0.80} \rfloor = 147$	0.789	0.808	0.845	0.820	0.809	0.803
m_{LR}	0.639	0.666	0.717	0.679	0.663	0.659
m	$t_{m,L}(\delta_0)$	$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\tau^*)$		
$\lfloor T^{0.40} \rfloor = 12$	0.006	0.010	0.013	0.009	0.009	0.010
$\lfloor T^{0.45} \rfloor = 16$	0.024	0.036	0.046	0.032	0.032	0.034
$\lfloor T^{0.50} \rfloor = 22$	0.065	0.085	0.108	0.080	0.079	0.081
$\lfloor T^{0.55} \rfloor = 30$	0.126	0.156	0.201	0.153	0.150	0.153
$\lfloor T^{0.60} \rfloor = 42$	0.243	0.281	0.343	0.280	0.273	0.276
$\lfloor T^{0.65} \rfloor = 57$	0.396	0.436	0.504	0.440	0.425	0.428
$\lfloor T^{0.70} \rfloor = 78$	0.577	0.611	0.672	0.623	0.607	0.604
$\lfloor T^{0.75} \rfloor = 107$	0.761	0.786	0.828	0.798	0.784	0.780
$\lfloor T^{0.80} \rfloor = 147$	0.895	0.907	0.927	0.914	0.907	0.904
m_{LR}	0.796	0.817	0.853	0.826	0.814	0.813

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$.

Table S.12: Empirical power. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = -0.15$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0$. $T = 1024$. Nominal asymptotic 5% level tests.

β_2	0		0	0.5	1	2
m	$LM_m(\delta_0)$	$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.005	0.006	0.008	0.006	0.006	0.006
$\lfloor T^{0.45} \rfloor = 16$	0.015	0.020	0.027	0.021	0.020	0.020
$\lfloor T^{0.50} \rfloor = 22$	0.052	0.067	0.086	0.064	0.065	0.066
$\lfloor T^{0.55} \rfloor = 30$	0.130	0.156	0.195	0.148	0.151	0.154
$\lfloor T^{0.60} \rfloor = 42$	0.263	0.295	0.354	0.291	0.286	0.293
$\lfloor T^{0.65} \rfloor = 57$	0.489	0.524	0.583	0.524	0.516	0.517
$\lfloor T^{0.70} \rfloor = 78$	0.731	0.758	0.798	0.761	0.751	0.752
$\lfloor T^{0.75} \rfloor = 107$	0.913	0.922	0.940	0.924	0.919	0.919
$\lfloor T^{0.80} \rfloor = 147$	0.986	0.987	0.990	0.988	0.987	0.987
m_{LR}	0.944	0.949	0.961	0.951	0.948	0.947
m	$t_{m,L}(\delta_0)$	$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\tau^*)$		
$\lfloor T^{0.40} \rfloor = 12$	0.022	0.032	0.041	0.030	0.031	0.032
$\lfloor T^{0.45} \rfloor = 16$	0.064	0.082	0.104	0.078	0.079	0.083
$\lfloor T^{0.50} \rfloor = 22$	0.150	0.181	0.222	0.178	0.176	0.178
$\lfloor T^{0.55} \rfloor = 30$	0.276	0.314	0.377	0.307	0.307	0.311
$\lfloor T^{0.60} \rfloor = 42$	0.453	0.492	0.559	0.484	0.481	0.487
$\lfloor T^{0.65} \rfloor = 57$	0.679	0.705	0.756	0.705	0.699	0.700
$\lfloor T^{0.70} \rfloor = 78$	0.863	0.877	0.904	0.878	0.872	0.874
$\lfloor T^{0.75} \rfloor = 107$	0.966	0.971	0.978	0.971	0.970	0.969
$\lfloor T^{0.80} \rfloor = 147$	0.995	0.996	0.997	0.996	0.996	0.995
m_{LR}	0.978	0.982	0.987	0.982	0.981	0.981

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$.

Table S.13: Empirical power. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = 0.15$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0$. $T = 512$. Nominal asymptotic 5% level tests.

β_2	0		0	0.5	1	2
m	$LM_m(\delta_0)$	$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40}\rfloor = 12$	0.054	0.017	0.004	0.005	0.012	0.017
$\lfloor T^{0.45}\rfloor = 16$	0.101	0.038	0.012	0.018	0.029	0.036
$\lfloor T^{0.50}\rfloor = 22$	0.169	0.079	0.024	0.041	0.063	0.075
$\lfloor T^{0.55}\rfloor = 30$	0.257	0.142	0.055	0.082	0.118	0.136
$\lfloor T^{0.60}\rfloor = 42$	0.384	0.248	0.119	0.165	0.214	0.240
$\lfloor T^{0.65}\rfloor = 57$	0.526	0.386	0.219	0.278	0.344	0.376
$\lfloor T^{0.70}\rfloor = 78$	0.669	0.548	0.365	0.433	0.502	0.539
$\lfloor T^{0.75}\rfloor = 107$	0.804	0.713	0.550	0.615	0.677	0.703
$\lfloor T^{0.80}\rfloor = 147$	0.906	0.851	0.736	0.784	0.828	0.843
m_{LR}	0.699	0.585	0.421	0.484	0.549	0.577
m	$t_{m,R}(\delta_0)$	$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\tau^*)$		
$\lfloor T^{0.40}\rfloor = 12$	0.100	0.036	0.013	0.015	0.026	0.034
$\lfloor T^{0.45}\rfloor = 16$	0.161	0.070	0.023	0.037	0.058	0.069
$\lfloor T^{0.50}\rfloor = 22$	0.244	0.129	0.050	0.075	0.108	0.124
$\lfloor T^{0.55}\rfloor = 30$	0.351	0.212	0.095	0.136	0.182	0.206
$\lfloor T^{0.60}\rfloor = 42$	0.488	0.341	0.183	0.240	0.302	0.330
$\lfloor T^{0.65}\rfloor = 57$	0.624	0.489	0.310	0.376	0.445	0.479
$\lfloor T^{0.70}\rfloor = 78$	0.750	0.645	0.472	0.538	0.606	0.636
$\lfloor T^{0.75}\rfloor = 107$	0.866	0.792	0.652	0.709	0.761	0.785
$\lfloor T^{0.80}\rfloor = 147$	0.940	0.904	0.816	0.854	0.886	0.899
m_{LR}	0.783	0.684	0.536	0.593	0.648	0.677
M	$KPSS$	$KPSS(\tau^*)$		$KPSS(\widehat{\tau})$		
M_e	0.293	0.313	0.193	0.206	0.272	0.304
M_{LR}	0.189	0.179	0.118	0.114	0.152	0.173

Note: $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.

Table S.14: Empirical power. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = 0.15$, $\varepsilon_t \sim NIID(0,1)$. Tests of $H_0 : \delta = 0$. $T = 1024$. Nominal asymptotic 5% level tests.

β_2	0		0	0.5	1	2
m	$LM_m(\delta_0)$	$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.098	0.038	0.011	0.019	0.030	0.037
$\lfloor T^{0.45} \rfloor = 16$	0.165	0.075	0.027	0.047	0.063	0.071
$\lfloor T^{0.50} \rfloor = 22$	0.270	0.151	0.064	0.103	0.135	0.147
$\lfloor T^{0.55} \rfloor = 30$	0.409	0.266	0.134	0.194	0.244	0.262
$\lfloor T^{0.60} \rfloor = 42$	0.572	0.430	0.258	0.341	0.403	0.423
$\lfloor T^{0.65} \rfloor = 57$	0.740	0.628	0.456	0.547	0.605	0.624
$\lfloor T^{0.70} \rfloor = 78$	0.875	0.809	0.681	0.746	0.793	0.804
$\lfloor T^{0.75} \rfloor = 107$	0.962	0.936	0.870	0.908	0.927	0.935
$\lfloor T^{0.80} \rfloor = 147$	0.994	0.987	0.968	0.979	0.985	0.987
m_{LR}	0.919	0.873	0.775	0.830	0.861	0.869
m	$t_{m,R}(\delta_0)$	$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\tau^*)$		
$\lfloor T^{0.40} \rfloor = 12$	0.158	0.068	0.023	0.041	0.058	0.066
$\lfloor T^{0.45} \rfloor = 16$	0.240	0.123	0.049	0.083	0.109	0.120
$\lfloor T^{0.50} \rfloor = 22$	0.368	0.227	0.107	0.161	0.203	0.222
$\lfloor T^{0.55} \rfloor = 30$	0.508	0.365	0.204	0.279	0.340	0.358
$\lfloor T^{0.60} \rfloor = 42$	0.667	0.537	0.357	0.448	0.510	0.533
$\lfloor T^{0.65} \rfloor = 57$	0.813	0.723	0.567	0.651	0.701	0.717
$\lfloor T^{0.70} \rfloor = 78$	0.918	0.869	0.769	0.823	0.856	0.865
$\lfloor T^{0.75} \rfloor = 107$	0.978	0.961	0.919	0.942	0.956	0.960
$\lfloor T^{0.80} \rfloor = 147$	0.997	0.994	0.983	0.990	0.993	0.994
m_{LR}	0.953	0.921	0.850	0.890	0.914	0.918
M	$KPSS$	$KPSS(\tau^*)$		$KPSS(\widehat{\tau})$		
M_e	0.345	0.382	0.252	0.288	0.358	0.379
M_{LR}	0.226	0.235	0.152	0.165	0.212	0.233

Note: $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.

Table S.15: Empirical power. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = -0.15$, $\varepsilon_t = \sigma_t z_t$, $z_t \sim NIID(0,1)$, $\sigma_t^2 = 1/(1-0.5) + 0.5\varepsilon_{t-1}^2$. Tests of $H_0 : \delta = 0$. $T = 512$. Nominal asymptotic 5% level tests.

β_2	0		0	0.5	1	2
m	$LM_m(\delta_0)$	$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.001	0.001	0.001	0.001	0.001	0.001
$\lfloor T^{0.45} \rfloor = 16$	0.003	0.005	0.005	0.005	0.004	0.004
$\lfloor T^{0.50} \rfloor = 22$	0.015	0.020	0.024	0.019	0.019	0.020
$\lfloor T^{0.55} \rfloor = 30$	0.041	0.052	0.063	0.050	0.049	0.052
$\lfloor T^{0.60} \rfloor = 42$	0.109	0.132	0.156	0.129	0.126	0.130
$\lfloor T^{0.65} \rfloor = 57$	0.230	0.260	0.300	0.262	0.254	0.256
$\lfloor T^{0.70} \rfloor = 78$	0.403	0.434	0.480	0.436	0.429	0.427
$\lfloor T^{0.75} \rfloor = 107$	0.593	0.615	0.661	0.624	0.616	0.611
$\lfloor T^{0.80} \rfloor = 147$	0.747	0.761	0.788	0.770	0.764	0.759
m_{LR}	0.614	0.635	0.670	0.639	0.632	0.631
m	$t_{m,L}(\delta_0)$	$t_{m,L}(\delta_0;\tau^*)$		$t_{m,L}(\delta_0;\tau^*)$		
$\lfloor T^{0.40} \rfloor = 12$	0.006	0.010	0.011	0.011	0.010	0.010
$\lfloor T^{0.45} \rfloor = 16$	0.024	0.034	0.039	0.031	0.032	0.033
$\lfloor T^{0.50} \rfloor = 22$	0.063	0.086	0.098	0.080	0.082	0.083
$\lfloor T^{0.55} \rfloor = 30$	0.123	0.155	0.179	0.148	0.150	0.151
$\lfloor T^{0.60} \rfloor = 42$	0.244	0.283	0.321	0.276	0.272	0.278
$\lfloor T^{0.65} \rfloor = 57$	0.402	0.444	0.492	0.442	0.433	0.436
$\lfloor T^{0.70} \rfloor = 78$	0.579	0.611	0.656	0.611	0.602	0.602
$\lfloor T^{0.75} \rfloor = 107$	0.738	0.757	0.791	0.762	0.754	0.753
$\lfloor T^{0.80} \rfloor = 147$	0.839	0.851	0.870	0.856	0.852	0.850
m_{LR}	0.747	0.768	0.794	0.773	0.767	0.762

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$.

Table S.16: Empirical power. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = -0.15$, $\varepsilon_t = \sigma_t z_t$, $z_t \sim NIID(0,1)$, $\sigma_t^2 = 1/(1-0.5) + 0.5\varepsilon_{t-1}^2$. Tests of $H_0 : \delta = 0$. $T = 1024$. Nominal asymptotic 5% level tests.

β_2	0		0	0.5	1	2
m	$LM_m(\delta_0)$	$LM_m(\delta_0;\tau^*)$	$LM_m(\delta_0;\widehat{\tau})$			
$\lfloor T^{0.40} \rfloor = 12$	0.003	0.006	0.007	0.005	0.006	0.006
$\lfloor T^{0.45} \rfloor = 16$	0.012	0.016	0.021	0.016	0.016	0.016
$\lfloor T^{0.50} \rfloor = 22$	0.052	0.066	0.080	0.063	0.062	0.065
$\lfloor T^{0.55} \rfloor = 30$	0.129	0.153	0.182	0.145	0.148	0.152
$\lfloor T^{0.60} \rfloor = 42$	0.265	0.296	0.334	0.290	0.288	0.291
$\lfloor T^{0.65} \rfloor = 57$	0.496	0.531	0.575	0.523	0.519	0.525
$\lfloor T^{0.70} \rfloor = 78$	0.728	0.750	0.781	0.746	0.738	0.743
$\lfloor T^{0.75} \rfloor = 107$	0.886	0.895	0.909	0.896	0.892	0.891
$\lfloor T^{0.80} \rfloor = 147$	0.952	0.956	0.961	0.956	0.955	0.954
m_{LR}	0.907	0.914	0.927	0.914	0.912	0.911
m	$t_{m,L}(\delta_0)$	$t_{m,L}(\delta_0;\tau^*)$	$t_{m,L}(\delta_0;\tau^*)$			
$\lfloor T^{0.40} \rfloor = 12$	0.022	0.032	0.037	0.030	0.032	0.031
$\lfloor T^{0.45} \rfloor = 16$	0.059	0.081	0.091	0.076	0.077	0.081
$\lfloor T^{0.50} \rfloor = 22$	0.143	0.175	0.204	0.167	0.171	0.173
$\lfloor T^{0.55} \rfloor = 30$	0.276	0.312	0.352	0.301	0.302	0.310
$\lfloor T^{0.60} \rfloor = 42$	0.447	0.488	0.538	0.479	0.477	0.486
$\lfloor T^{0.65} \rfloor = 57$	0.680	0.710	0.749	0.703	0.699	0.704
$\lfloor T^{0.70} \rfloor = 78$	0.852	0.866	0.888	0.864	0.861	0.861
$\lfloor T^{0.75} \rfloor = 107$	0.941	0.946	0.955	0.947	0.944	0.944
$\lfloor T^{0.80} \rfloor = 147$	0.974	0.975	0.978	0.976	0.975	0.975
m_{LR}	0.953	0.957	0.963	0.957	0.955	0.954

Note: $t_{m,L}$ denotes alternative $H_1 : \delta < 0$.

Table S.17: Empirical power. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = 0.15$, $\varepsilon_t = \sigma_t z_t$, $z_t \sim NIID(0,1)$, $\sigma_t^2 = 1/(1-0.5) + 0.5\varepsilon_{t-1}^2$. Tests of $H_0 : \delta = 0$. $T = 512$. Nominal asymptotic 5% level tests.

β_2	0		0	0.5	1	2
m	$LM_m(\delta_0)$	$LM_m(\delta_0;\tau^*)$	$LM_m(\delta_0;\widehat{\tau})$			
$\lfloor T^{0.40} \rfloor = 12$	0.049	0.017	0.005	0.006	0.012	0.016
$\lfloor T^{0.45} \rfloor = 16$	0.097	0.037	0.011	0.018	0.029	0.035
$\lfloor T^{0.50} \rfloor = 22$	0.165	0.074	0.023	0.042	0.060	0.071
$\lfloor T^{0.55} \rfloor = 30$	0.261	0.138	0.056	0.084	0.118	0.132
$\lfloor T^{0.60} \rfloor = 42$	0.381	0.247	0.121	0.170	0.214	0.237
$\lfloor T^{0.65} \rfloor = 57$	0.522	0.382	0.222	0.287	0.343	0.371
$\lfloor T^{0.70} \rfloor = 78$	0.659	0.540	0.375	0.439	0.502	0.530
$\lfloor T^{0.75} \rfloor = 107$	0.779	0.690	0.543	0.603	0.656	0.680
$\lfloor T^{0.80} \rfloor = 147$	0.864	0.805	0.697	0.743	0.783	0.798
m_{LR}	0.670	0.554	0.400	0.466	0.524	0.546
m	$t_{m,R}(\delta_0)$	$t_{m,R}(\delta_0;\tau^*)$	$t_{m,R}(\delta_0;\tau^*)$			
$\lfloor T^{0.40} \rfloor = 12$	0.097	0.034	0.011	0.017	0.026	0.032
$\lfloor T^{0.45} \rfloor = 16$	0.161	0.070	0.022	0.036	0.057	0.067
$\lfloor T^{0.50} \rfloor = 22$	0.241	0.128	0.051	0.076	0.108	0.123
$\lfloor T^{0.55} \rfloor = 30$	0.355	0.215	0.098	0.137	0.184	0.205
$\lfloor T^{0.60} \rfloor = 42$	0.486	0.339	0.185	0.249	0.302	0.328
$\lfloor T^{0.65} \rfloor = 57$	0.616	0.487	0.314	0.383	0.448	0.477
$\lfloor T^{0.70} \rfloor = 78$	0.741	0.635	0.476	0.538	0.602	0.625
$\lfloor T^{0.75} \rfloor = 107$	0.841	0.768	0.641	0.694	0.741	0.759
$\lfloor T^{0.80} \rfloor = 147$	0.902	0.860	0.773	0.809	0.842	0.853
m_{LR}	0.752	0.649	0.508	0.567	0.620	0.642
M	$KPSS$	$KPSS(\tau^*)$	$KPSS(\widehat{\tau})$			
M_e	0.293	0.313	0.193	0.206	0.272	0.304
M_{LR}	0.188	0.174	0.123	0.119	0.151	0.168

Note: $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.

Table S.18: Empirical power. DGP $u_t = \Delta^{-\delta}\varepsilon_t$, $\delta = -0.15$, $\varepsilon_t = \sigma_t z_t$, $z_t \sim NIID(0,1)$, $\sigma_t^2 = 1/(1-0.5) + 0.5\varepsilon_{t-1}^2$. Tests of $H_0 : \delta = 0$. $T = 1024$. Nominal asymptotic 5% level tests.

β_2	0		0	0.5	1	2
m	$LM_m(\delta_0)$	$LM_m(\delta_0;\tau^*)$		$LM_m(\delta_0;\widehat{\tau})$		
$\lfloor T^{0.40} \rfloor = 12$	0.094	0.036	0.010	0.020	0.028	0.035
$\lfloor T^{0.45} \rfloor = 16$	0.160	0.074	0.024	0.046	0.062	0.071
$\lfloor T^{0.50} \rfloor = 22$	0.268	0.151	0.063	0.106	0.132	0.146
$\lfloor T^{0.55} \rfloor = 30$	0.402	0.266	0.135	0.198	0.240	0.260
$\lfloor T^{0.60} \rfloor = 42$	0.568	0.428	0.263	0.348	0.401	0.422
$\lfloor T^{0.65} \rfloor = 57$	0.735	0.627	0.452	0.546	0.603	0.619
$\lfloor T^{0.70} \rfloor = 78$	0.866	0.798	0.670	0.741	0.780	0.792
$\lfloor T^{0.75} \rfloor = 107$	0.947	0.918	0.847	0.887	0.908	0.915
$\lfloor T^{0.80} \rfloor = 147$	0.979	0.967	0.939	0.954	0.964	0.966
m_{LR}	0.906	0.856	0.759	0.811	0.843	0.851
m	$t_{m,R}(\delta_0)$	$t_{m,R}(\delta_0;\tau^*)$		$t_{m,R}(\delta_0;\tau^*)$		
$\lfloor T^{0.40} \rfloor = 12$	0.156	0.065	0.020	0.043	0.054	0.064
$\lfloor T^{0.45} \rfloor = 16$	0.236	0.120	0.047	0.082	0.105	0.117
$\lfloor T^{0.50} \rfloor = 22$	0.360	0.222	0.107	0.163	0.201	0.219
$\lfloor T^{0.55} \rfloor = 30$	0.508	0.363	0.206	0.281	0.332	0.356
$\lfloor T^{0.60} \rfloor = 42$	0.664	0.535	0.357	0.455	0.505	0.527
$\lfloor T^{0.65} \rfloor = 57$	0.810	0.723	0.568	0.651	0.702	0.717
$\lfloor T^{0.70} \rfloor = 78$	0.911	0.860	0.760	0.812	0.845	0.855
$\lfloor T^{0.75} \rfloor = 107$	0.964	0.945	0.896	0.926	0.939	0.944
$\lfloor T^{0.80} \rfloor = 147$	0.988	0.980	0.960	0.971	0.976	0.979
m_{LR}	0.939	0.903	0.833	0.875	0.895	0.902
M	$KPSS$	$KPSS(\tau^*)$		$KPSS(\widehat{\tau})$		
M_e	0.345	0.382	0.252	0.288	0.358	0.379
M_{LR}	0.230	0.231	0.159	0.167	0.208	0.228

Note: $t_{m,R}$ denotes alternative $H_1 : \delta > 0$.