Locally Optimal Invariant Tests for Perturbed Fractional Integration

Robert Taylor

University of Essex

Based on joint work with James Duffy (Oxford), David Harris (Melbourne) and Morten Nielsen (Aarhus)

Aarhus Workshop in Econometrics VI
November 28th, 2024

Outline of Presentation

Table of Contents

Introduction

- We develop a class of invariant tests for the null hypothesis that a time series is a martingale difference sequence against the alternative that it belongs to the class of perturbed fractionally integrated (long memory) processes.
- The tests are indexed by a user-chosen long memory parameter, $d > 0$, and are locally most powerful (under Gaussianity) where the true long memory parameter, d^* say, coincides with this value.
- d^* is a nuisance parameter, present only under the alternative.
- The class of tests contains a number of widely used tests as special cases, including the Nyblom and Mäkeläinen (1983) and Kwiatkowski *et al* (1992) [KPSS] tests.
- A taxonomy of limiting null distribution theory for the class of statistics (indexed by d) is provided. These null distributions depend on d.
- We compare the local power properties of the tests under appropriate Pitman drift sequences.
- Extensions to allow for the presence of a general form of weak dependence under the null hypothesis and deterministic mean components are also considered.

Table of Contents

The Perturbed Long Memory Model

- Consider the signal-plus-noise model:

$$
\begin{aligned}
y_t &= \mu_t + \varepsilon_t, \quad t = 1, ..., T && (1) \\
\mu_t &= \sum_{j=0}^{t-1} b_{d,j} \eta_t, && (2)
\end{aligned}
$$

where the weights

$$
b_{d,0} = 0, \quad \text{and} \quad b_{d,j} := \frac{\Gamma(d+j)}{\Gamma(d)\Gamma(1+j)} = \frac{d(d+1)\dots(d+j-1)}{j!}, \quad j = 1, 2, \dots \quad (3)
$$

are the coefficients in the usual binomial expansion of $(1-z)^{-d}$, $d \in \mathbb{R}^{+}$.

- In the context of (1)-(2), ε_t and η_t are assumed to be mutually independent zero mean weakly stationary, $I(0)$, short memory processes with (finite) variances σ_ε^2 and $\omega^2 \sigma_\varepsilon^2$, respectively, with $\sigma_\varepsilon^2 > 0$ and $\omega \geq 0$.

The Perturbed Long Memory Model

- For the set of weights in (3), we can equivalently write

$$\mu_t = \Delta_+^{-d} \eta_t$$

where, for a generic variable x_t,

$$\Delta_+^{-d} x_t := (1 - L)^{-d} x_t . \mathbb{I}(t \geq 1)$$

in which L is the usual lag operator such that $L^k x_t = x_{t-k}$, $k = 0, \pm 1, \pm 2, ...$, and $\mathbb{I}(\cdot)$ denotes the indicator function.

- We therefore have that μ_t is a (long memory) type II fractionally integrated process of order d, denoted $I(d)$; see, for example, Marinucci and Robinson (1999). As a consequence, y_t is also $I(d)$.

- One interpretation of the model is that of an $I(d)$ series, μ_t, which is observed subject to an $I(0)$ measurement error, ε_t; that is, a signal-plus-noise model, where the $I(d)$ signal process (or long-run component), μ_t, is perturbed by the additive $I(0)$ noise (or short run component), ε_t.

- Additive trend-cycle decomposition models of these forms have been widely applied in many fields, particularly in economics and finance.
- A notable special case of (1)-(2) is the local level (or random walk plus noise) unobserved components model of eg Harvey (1989).
- This model obtains on setting $d = 1$ in (2), with ε_t and η_t both IID, such that μ_t follows a random walk, and has a reduced form $ARIMA(0, 1, 1)$ representation (where the MA coefficient is non-positive) when ε_t and η_t are both white noise processes.
- The 'smooth trend' unobserved components model of Clark (1987), see also Harvey (1989), is a further special case, obtained for $d = 2$ with ε_t and η_t again both IID, which has a reduced form $ARIMA(0, 2, 2)$ representation.
- Where d is fractional, (1)-(2) does not have a reduced form ARFIMA representation except in the case where η_t is uncorrelated and ε_t admits an ARMA representation in the fractional lag operator (see Johansen, 2008, Equation 2).

- The local level and smooth trend models can also be interpreted as regression models in which the constant term evolves as a random walk and as an integrated random walk, respectively.
- The fractional long-run component in (1)-(2) allows considerable flexibility on the weighting on past shocks compared to eg the local level and smooth trend models.
- The parameter d determines the rate of decay of the autocovariance function of μ_t, and hence of y_t. For $0 < d < 0.5$, μ_t is (asymptotically) stationary. For $d < 1$, μ_t is mean-reverting while for $d > 1$, μ_t is a cumulated mean-reverting process.

The Perturbed Long Memory Model

- The fractional plus noise decomposition in (1)-(2) has proved particularly useful in modelling financial volatility.
- Of particular note is the long memory stochastic volatility [LMSV] model for financial returns considered in Breidt *et al.* (1998), Harvey (1998) and Bollerslev and Jubinski (1999).
- Here, returns, r_t, are modelled as $r_t = \kappa \exp\{(\mu_t + v_t)/2\} u_t$, where μ_t and v_t are, respectively, long and short memory processes, and μ_t, v_t and u_t are independent. For the LMSV model, r_t is a martingale difference sequence [MDS] if u_t is a MDS.
- Taking logs of the squared returns we then obtain that

$$\log r_t^2 = \mu_t + v_t + \log \kappa^2 + \log u_t^2$$

which corresponds to a fractional plus noise model of the form in (1)-(2) for $\log r_t^2$ where the signal μ_t corresponds to (the long memory component of) the log-volatility of returns and where the additive noise component is given by $\varepsilon_t = v_t + \log \kappa^2 + \log u_t^2$ which is, in general, assumed to be an IID process.

- For integer d, estimation of (1)-(2) can either be conduced using a state space representation of the model and an application of the Kalman filter algorithm, or by estimating the reduced form of the model using (restricted) quasi MLE.
- For the fractional case, estimation is considerably more complicated with conventional ARFIMA-based estimation not, in general, possible.
- A number of approaches have been considered which depend on the true long memory parameter and the specific assumptions made about ε_t and η_t
- Frederiksen *et al.* (2012) develop a semiparametric local polynomial Whittle estimator of d which they show to be consistent for $0 < d < 1$. Unlike most of the literature which assumes that ε_t is IID, Frederiksen *et al.* (2012) allow for weak dependence in ε_t.
- Applying this estimator to daily log-squared return series for the 30 Dow Jones Industrial Average [DJIA] stocks, Frederiksen *et al.* (2012) find that most of the stocks yield estimates of d in the nonstationary ($d > 1/2$) region.

Table of Contents

Extant LBI tests for $d = 1$ and $d = 2$

- For the local level model, testing for the presence of the $I(1)$ stochastic trend component, μ_t, has been widely studied.
- These can be interpreted either as a test for the null that the series is $I(0)$ against the alternative that it is $I(1)$, or as a test for the null of a constant level against the alternative that the level of the series evolves as a random walk.
- The most powerful invariant (likelihood ratio) test is a function of the true (unknown) value of signal-to-noise ratio, ω^2 - no uniformly most powerful test exists.
- Nyblom and Mäkeläinen (1983) and Nyblom (1986), among others, develop an LBI test of the null hypothesis that $\omega^2 = 0$, so that $\mu_t = 0$ for all t, in the case where ε_t and η_t are independent Gaussian white noises.
- Kwiatkowski *et al* (1992) modify these LBI tests to allow the random walk to be embedded within a more general weakly dependent process.
- Nyblom and Harvey (2001) develop an LBI test for the $d = 2$ case, testing the null that the series is $I(0)$ against the alternative that it is $I(2)$.
- LBI tests have the considerable advantage over likelihood-ratio based tests in this testing problem in that, being one-sided Lagrange Multiplier [LM] tests, they avoid the need to estimate the model under the alternative hypothesis.

Our Contribution

- We generalise the LBI testing approach developed in the context of the local level and smooth trend models to enable testing for the presence of the long memory signal, μ_t, in (1)-(2) for any user-specified value of d.
- The LBI testing principle only requires estimation of (1)-(2) under the null and so is particularly convenient to use here given the inherent problems with estimating (1)-(2) when d is fractional, noted above.
- Our tests are therefore of the null hypothesis that a series is $I(0)$ (either serially uncorrelated or weakly stationary, depending on the assumptions made on ε_t), against the alternative that the series is a perturbed fractionally integrated process.
- For any d such that $0 < d \leq 1/2$, we show that a suitably scaled version of the LBI statistic has a standard normal limiting null distribution, while for $d > 0.25$, the scaled LBI statistic admits a non-standard limiting null distribution whose functional form explicitly depends on d.
- Where $d > 1/2$ this distribution is the so-called Cramér-von Mises function of a standard fractional Brownian motion.
- Extensions to allow for the presence of weak dependence in ε_t, and deterministic mean components are considered.

Table of Contents

- In the context of (1)-(2) we are looking to test

$$\begin{aligned} H_0 : \omega^2 &= 0 \qquad \text{vs.} && (4) \\ H_1 : \omega^2 &> 0 && (5) \end{aligned}$$

where ω^2 is the signal-to-noise ratio of the fractional component, μ_t.

- Writing $\boldsymbol{\mu} := (\mu_1, \ldots, \mu_T)'$ and $\boldsymbol{\eta} := (\eta_1, \ldots, \eta_T)'$, we have that, for a given value of $d > 0$, $\boldsymbol{\mu} = \boldsymbol{C}_d \boldsymbol{\eta}$, where

$$\boldsymbol{C}_d := \begin{pmatrix} b_{d,0} & 0 & & & 0 \\ b_{d,1} & b_{d,0} & 0 & & \\ b_{d,2} & b_{d,1} & b_{d,0} & 0 & \\ \vdots & & & \ddots & 0 \\ b_{d,T-1} & & & & b_{d,0} \end{pmatrix}.$$

- To derive locally optimal tests of H_0 against H_1 we will assume that ε_t and η_t are independent Gaussian white noise processes. The assumption of Gaussianity facilitates the construction of LBI tests based on statistics with a simple structure. All of the limiting results we give for these statistics hold under a considerably weaker MDS assumption.
- Under the Gaussianity assumption, it is straightforwardly seen from (1)-(2) that

$$\boldsymbol{y} := (y_1, \ldots, y_T)' \sim N_T\left(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{C}_d(\omega^2)\right), \tag{6}$$

where N_k denotes a multivariate Gaussian distribution of dimension k, and

$$\boldsymbol{C}_d(\omega^2) := \boldsymbol{I}_T + \omega^2 \boldsymbol{C}_d \boldsymbol{C}_d'.$$

- From (6), the testing problem (4),(5) is seen to be invariant under the group of transformations $\boldsymbol{y} \to \gamma_0 \boldsymbol{y}$ and $\sigma_\varepsilon^2 \to \gamma_0 \sigma_\varepsilon^2$, where γ_0 is a positive scalar. A maximal invariant for this testing problem is given by $\mathbf{w} := (\boldsymbol{y}'\boldsymbol{y})^{-1/2}\boldsymbol{y}$.

LBI Tests for Arbitrary d

- Following King and Hillier (1985,p.99), an application of the Neyman-Pearson Lemma yields that a most powerful invariant test for H_0 against the point alternative $\omega^2 = \omega_1^2 > 0$ in the context of (1)-(2), is defined by the critical region

$$\mathcal{LR} := \mathbf{w}'(\boldsymbol{C}_d(\omega_1^2))^{-1}\mathbf{w} < \ell, \tag{7}$$

  for some constant ℓ.

- While $\mathcal{LR}$ is useful for determining the (exact) Gaussian power envelope for the testing problem (4),(5), the functional form of $\mathcal{LR}$ is a function of ω_1^2 and, hence, no uniformly most powerful test exists for this testing problem.

LBI Tests for Arbitrary d

- Noting that $\boldsymbol{C}_d(0) = \boldsymbol{I}_T$, it follows from King and Hiller (1985,p.99) that an LBI test exists for the testing problem (4), (5).

- An LBI test is defined as a test which, for a given testing problem, has a power function which has maximum slope at the origin among all invariant tests.

- From Ferguson (1967,p.235) an LBI test is defined by the critical region

$$\left.\frac{\partial \ln f(\mathbf{w}|\omega^2)}{\partial \omega^2}\right|_{\omega^2=0} > \ell^* \tag{8}$$

for some suitably chosen constant ℓ^*, where, aside from irrelevant constants,

$$\ln f(\mathbf{w}|\omega^2) := -\frac{T}{2} \log \mathbf{w}' \boldsymbol{C}_d(\omega^2)^{-1} \mathbf{w} \tag{9}$$

is the natural logarithm of the likelihood function of $\mathbf{w}$.

- Consequently, using an application of King and Hiller (1985, Eqn.(6), p.99), this implies that an LBI test of H_0 against H_1 rejects for large values of the statistic

$$S_T(d) := \frac{\boldsymbol{y}'\boldsymbol{C}_d\boldsymbol{C}_d'\boldsymbol{y}}{\boldsymbol{y}'\boldsymbol{y}} = \frac{\sum_{t=1}^{T}\left(\sum_{i=t}^{T} b_{d,i-t}y_i\right)^2}{\sum_{t=1}^{T} y_t^2}. \tag{10}$$

- For $d = 1$, the weights are such that $b_{1,j} = 1$, $j = 0, 1, 2...$, and $S_T(1)$ coincides with the statistic L. Rejecting for large values of L in this setting is a LBI test for the null that y_t is a Gaussian white noise process against the alternative that it follows a Gaussian random walk plus noise model.
- Setting $d = 2$ in (14), $S_T(2)$ coincides with the ℓ_{IRW} statistic (albeit applied to $\boldsymbol{y}$ rather than to regression residuals) of Nyblom and Harvey (2001) which is LBI for testing the null that y_t is a Gaussian white noise process against the alternative that it follows a Gaussian integrated random walk plus noise model.
- The LBI test in (14) generalises these to allow testing the null that the series is a Gaussian white noise process against the alternative that it follows a Gaussian fractionally integrated, $I(d)$, plus noise model.

- DGP (2) defines μ_t to follow a type II fractionally integrated process. In contrast, a type I fractionally integrated process takes the form $(1-L)^d \mu_t = \eta_t$, giving

$$\mu_t = \sum_{j=0}^{\infty} b_{d,j} \eta_{t-j}.$$

- As is well known, this model is only well defined for $d < 1/2$. So although an LBI test of H_0 versus H_1 could potentially be developed, it would be restricted to stationary values of d. We therefore prefer the type II model for μ_t because it allows us to test any $d > 0$, rather than restricting d to be in the range $(0, 1/2)$.

- That said, the LBI test designed for the type II case would likely still have power against H_1 in the type I case.

Two Discussion Points: 2. Correlated Components

- In common with the majority of the literature on unobserved components models, including on LMSV models, we have assumed that ε_t and η_t in (1)-(2) are independent.

- However, the unobserved components literature has considered models which allow for a contemporaneous correlation between the shocks - in our case ε_t and η_t.

- This obtains on setting $E(\epsilon_t, \eta_t) = \rho\omega\sigma_\varepsilon^2$, where $|\rho| \leq 1$; a special case thereof, where $\rho = 1$, is the single innovation form of (1)-(2) where $y_t = \mu_t + \varepsilon_t$ with μ_t generated by $\mu_t = \omega\Delta_+^{-d}\varepsilon_t$.

- Provided $\omega > 0$, the correlation between ϵ_t and η_t is measured by ρ and is scale independent of ω and σ_ε^2. Otherwise, ρ is not defined, the component μ_t vanishing from (2).

Two Discussion Points: 2. Correlated Components

- For this model, $\boldsymbol{y} \sim N_T(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{D}_d(\omega))$ where $\boldsymbol{D}_d(\omega) := \boldsymbol{C}_d(\omega^2) + \rho\omega(\boldsymbol{C}_d + \boldsymbol{C}_d')$. Hence, the density of the maximal invariant $\mathbf{w}$, is given by (9) but with $\boldsymbol{C}_d(\omega^2)$ replaced by $\boldsymbol{D}_d(\omega)$.
- As would be expected, (14) defines the LBI test of $H_0^* : \omega = 0$ against $H_1^* : \omega > 0$ when $\rho = 0$.
- When $\rho \neq 0$, an application of (8) yields that the LBI test of H_0^* against H_1^* rejects for large values of the statistic

$$S_T^*(d) := \frac{\boldsymbol{y}'\boldsymbol{C}_d\boldsymbol{y}}{\boldsymbol{y}'\boldsymbol{y}} = \frac{\sum_{t=1}^T y_t \left(\sum_{j=0}^{t-1} b_{d,j} y_{t-j}\right)}{\sum_{t=1}^T y_t^2} = \sum_{j=0}^{T-1} b_{d,j} \frac{\sum_{t=j+1}^T y_t y_{t-j}}{\sum_{t=1}^T y_t^2} = \sum_{j=0}^{T-1} b_{d,j} \hat{r}_j.$$

where $\hat{r}_j$ denotes the jth sample autocorrelation of y_t.

- Notice that the functional form of $S^*_T(d)$ does not depend on ρ.
- Noting that $b_{d,j} \sim j^{d-1}$, we have that

$$S^*_T(d) \simeq \sum_{j=1}^{T-1} j^{d-1} \hat{r}_j$$

.

- Setting $d = 0$, the left hand side of the approximation simplifies to the statistic upon which the LBI test (which is also a uniformly most powerful test) of Tanaka (1999) for testing $H_{0,d} : d = 0$ against $H_{1,d} : d > 0$ in the case where $y_t = \Delta_+^{-d} \eta_t$, with η_t a Gaussian white noise, is based; see Tanaka (1999, Eqn.(40), p.560).
- Consequently, choosing d arbitrarily close to zero, a test which rejects for large values of $S^*_T(d)$ has approximately the same critical region as the LBI test of $H_{0,d} : d = 0$ against $H_{1,d} : d > 0$ in the non-perturbed fractional model.

Table of Contents

Null Asymptotics

- We establish the limiting null distributions of the family of $S_T(d)$ statistics, indexed by d. As we will show, these limiting distribution fall into three different categories, depending on whether $d \in (0, 1/4]$, $d \in (1/4, 1/2]$, or $d \in (1/2, \infty)$.
- For each of these ranges different scalings and/or centerings of $S_T(d)$ are required to obtain a convergent distribution, free of nuisance parameters. In each case, these terms are deterministic and so the LBI property is preserved.
- In doing so, we can considerably relax the Gaussian IID assumption on $\{\varepsilon_t\}$ to the following MDS assumption which is conventional in the long memory literature.

Assumption 1

$\{\varepsilon_t, \mathcal{F}_t\}_{t\in\mathbb{Z}}$ is a stationary MDS with $E(\varepsilon_t^2 \mid \mathcal{F}_{t-1}) = \sigma_\varepsilon^2$ which is a.s. constant, and $\sup_{t\in\mathcal{N}} E(\varepsilon_t^4) < \infty$.

- We will discuss later how this could be further relaxed, at least for some values of d, to allow for weak dependence in ε_t.

- In particular, for $d \in (0, 1/2]$, it will prove convenient, in terms of developing a statistic with a limiting null distribution free of nuisance parameters, to develop an equivalent test based on the centred statistic.

$$S_{2,T}(d) := S_T(d) - \sum_{i=0}^{T-1} b_{d,i}^2.$$

- For $d \in (0, 1/4]$ we also need to define the quantity

$$\sigma_T^2 \quad := \quad \sum_{t=2}^{T} \sum_{s=1}^{t-1} a_T^2(s,t),$$

where

$$a_T(s,t) \quad := \quad \sum_{i=0}^{s-1} b_{d,i} b_{d,i+t-s}$$

- For $d \in (1/4, 1/2]$ the limiting null distribution will feature the limiting RV

$$\int_0^1 \int_0^r f_d(r,q)\mathrm{d}W(q)\mathrm{d}W(r) := \mathcal{I}(f_d, W)$$

where $W(\cdot)$ denotes a standard Brownian motion on $[0,1]$, and the function

$$f_d(r,q) := \begin{cases} \int_0^1 (x-r)_+^{d-1}(x-q)_+^{d-1}\mathrm{d}x & \text{if } r, q \in [0,1] \text{ with } q < r, \\ 0 & \text{otherwise;} \end{cases} \tag{11}$$

where $(a)_+ := \max\{a, 0\}$, and $d \in (0,1)$.

Null Asymptotics

- For $d \in (1/2, \infty)$ the limiting null distribution will feature the limiting RV

$$W_d(r) := \int_0^r (r-q)^{d-1} \mathrm{d}W(q). \tag{12}$$

- The process $W_d(\cdot)$ is the familiar type II fractional Brownian motion of order d.
- Observe that $W_1(r) = W(r)$.

Null Asymptotics

Theorem 1

Consider the statistic S_T of (14). Let y_t be generated according to (1)-(2) under H_0 of (4). Then under Assumption 1,

(i) *if* $d \in (0, 1/4]$, *then*

$$(T\sigma_T^{-1})S_{2,T}(d) \rightsquigarrow N[0,4];$$

(ii) *if* $d \in (1/4, 1/2]$, *then*

$$(Tb_{d,T}^2)^{-1}S_{2,T}(d) \rightsquigarrow 2\mathcal{I}(f_d, W) - (2d)^{-1};$$

(iii) *if* $d \in (1/2, \infty)$, *then*

$$(Tb_{d,T}^2)^{-1}S_T(d) \rightsquigarrow \int_0^1 W_d^2(r)\mathrm{d}r.$$

- The limiting distribution in part (iii) reduces for $d = 1$ and $d = 2$ to the limiting results given in Nyblom and Mäkeläinen (1983) and Nyblom and Harvey (2001), respectively.
- In the I(1) case the limiting null distribution $\int_0^1 W^2(r)\mathrm{d}r$ is the so-called Cramér-von Mises distribution of order 1, $CvM(1)$, which has been widely tabulated in the literature.
- In the I(2) case this is seen by noting that $W_2(r) = \int_0^r (r-q)\mathrm{d}W(q) \overset{a.s.}{=} \int_0^r W(q)\mathrm{d}q$, by an application of the (stochastic) Fubini Theorem, which is the integrated Brownian motion form of the process given in NH and is tabulated there.
- For $d > 1/2$ we will refer to the limiting distribution in part (iii) of Theorem 1 as a fractional Cramer-von Mises distribution, or $CvM(d)$.
- The limiting distribution in part (ii) appears to be new to the literature.

- The fixed alternative H_1 is that $\omega^2 > 0$, and hence that μ_t evolves as an $I(d)$ process.
- As with most testing problems, a more useful measure for assessing the relative power of competing tests is to look at local power under Pitman drift sequences.
- The Pitman drift rate needs to relate to the true long memory parameter, say d^*, which is a nuisance parameter only present under the alternative.
- The Pitman drift rate for our testing problem is of the form

$$\omega = \omega_{d^*,T} = \frac{c}{T d_{d^*,T}} \approx \frac{c}{T^{d^*}}$$

- Notice that this Pitman rate does not depend on the value of d used in constructing the LBI statistic, $S_T(d)$.

- Rather than detail the representations for these asymptotic local power functions (these are given in the paper) I will present simulations of the local power functions of the tests for selected values of d in the simulation results later.

- This will allow us to compare the local power of $S_T(d)$ for various choices of d and d^*. Recall that the LBI property holds only where $d = d^*$.

- Based on these results we will make some practical testing recommendations, given that d^* is not known in practice, but d can be chosen by the practitioner.

Table of Contents

Deterministic Components

- Replace (1) by

$$y_t = d_t + \mu_t + \varepsilon_t \tag{13}$$

  where μ_t and ε_t are as defined for (1) and where d_t is a deterministic kernel of the form $d_t = \mathbf{x}_t'\beta$ where $\mathbf{x}_t$ is a $(p \times 1)$, $p < T$, fixed sequence, whose first element is fixed at unity throughout (so that (13) contains an intercept term), with associated parameter vector β.

- In the context of (13), the testing problem (4)-(5) is invariant under the group of transformations $\mathbf{y} \to \gamma_0\mathbf{y} + \mathbf{X}\gamma$, $\beta \to \gamma_0\beta + \gamma$ and $\sigma_\varepsilon^2 \to \gamma_0^2\sigma_\varepsilon^2$, where $\mathbf{X} := [\mathbf{x}_1, \ldots, \mathbf{x}_T]'$, γ_0 is a positive scalar and γ a real $(k \times 1)$ vector.

- A maximal invariant for this hypothesis testing problem is now given by $\mathbf{w_x} := (\mathbf{e}'\mathbf{e})^{-1/2}\mathbf{Pe}$, where $\mathbf{P}$ is the $(T-k) \times T$ matrix such that $\mathbf{PP}' = \mathbf{I}_{T-k}$, $\mathbf{P}'\mathbf{P} = \mathbf{M}$, $\mathbf{M} := \mathbf{I}_T - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, and $\mathbf{e} := (\hat{e}_1, ..., \hat{e}_T)' = \mathbf{My}$ is the vector of OLS residuals from the regression of $\mathbf{y}$ on $\mathbf{X}$.

- Notice that $\{\hat{e}_t\}$ are the OLS residual from estimating (13) under H_0.
- Applying (8), we obtain that, under Gaussianity, the LBI test rejects for large values of the statistic

$$S_{T,\mathbf{x}}(d) := \frac{\boldsymbol{e}'\boldsymbol{C}_d\boldsymbol{C}_d'\boldsymbol{e}}{\boldsymbol{e}'\boldsymbol{e}} = \frac{\sum_{t=1}^{T}\left(\sum_{i=t}^{T} b_{d,i-t}\hat{e}_i\right)^2}{\sum_{t=1}^{T}\hat{e}_t^2}. \tag{14}$$

- In this setting, again using King and Hillier (1985,p.99), the decision rule for the MPI test for H_0 against the point alternative $\omega^2 = \omega_1^2 > 0$, again under the Gaussianity assumption, is defined by the critical region,

$$\mathcal{LR}_{\mathbf{x}} := \mathbf{w}_{\mathbf{x}}'(\mathbf{P}\boldsymbol{C}_d(\omega_1^2)\mathbf{P}')^{-1}\mathbf{w}_{\mathbf{x}} < \ell \tag{15}$$

for some constant ℓ, where $\mathbf{w}_{\mathbf{x}}$ and $\mathbf{P}$ are as defined in the previous paragraph.

- Assume that the vector $\mathbf{x}_t$ satisfies the following conditions (see Phillips and Xiao, 1998): there exists a scaling matrix δ_T and a bounded piecewise continuous function $\mathbf{x}(r)$ such that: (i) $\delta_T \mathbf{x}_{\lfloor Tr \rfloor} \to \mathbf{x}(r)$ as $T \to \infty$ uniformly in $r \in [0,1]$, and (ii) $\int_0^1 \mathbf{x}(r)\mathbf{x}(r)'\mathrm{d}r$ is positive definite.
- In this case for the $d = 1$ case, it is known that

$$T^{-1} S_{T,\mathbf{x}}(1) \Rightarrow \int_0^1 W_{\mathbf{x}}(r)^2 \mathrm{d}r \tag{16}$$

where

$$W_{\mathbf{x}}(r) := W(r) - \int_0^1 \mathbf{x}(r)'\mathrm{d}W(r) \left(\int_0^1 \mathbf{x}(r)\mathbf{x}(r)'\mathrm{d}r \right)^{-1} \int_0^r \mathbf{x}(s)\mathrm{d}s, \quad r \in [0,1] \tag{17}$$

where it is recalled that $W(\cdot)$ is a standard Brownian motion process.

- In the case where $\mathbf{x}_t = 1$, $W_{\mathbf{x}}(r) = W(r) - rW(1) =: B_0(r)$, $r \in [0,1]$, a standard Brownian bridge process, and the right member of (16) is a first level Cramér-von Mises distribution with one degree of freedom, denoted $CvM_1(1)$.
- Where $\mathbf{x}_t = (1, t)'$, $W_{\mathbf{x}}(r) := B_0(r) - 6r(1-r)\int_0^1 B_0(s)\mathrm{d}s$, $r \in [0,1]$, a standard second level Brownian bridge process.

Deterministic Components

- We are currently working on the corresponding limiting distributions in the fractional d case. This is complicated somewhat by the fact that the inner summation in the LBI statistics appears to only be reversable for integer values of d.
- So far, we have a complete set of limiting null distributions for the case where $\mathbf{x}_t = 1$. Define the version of $S_T(d)$ calculated from the residuals obtained under the null when $\mathbf{x}_t = 1$ as $S_T^{\mu}(d)$.
- In particular, the result in part (i) of Theorem 1 still holds for a centred version of $S_T^{\mu}(d)$; that is,

  $$(T\sigma_T^{-1})S_{2,T}^{\mu}(d) \rightsquigarrow N[0,4]$$

  where $S_{2,T}^{\mu}(d) := S_T^{\mu}(d) - \sum_{i=0}^{T-1} b_{d,i}^2$.
- For part (iii), the limiting null distribution changes:

  $$(Tb_{d,T}^2)^{-1}S_T^{\mu}(d) \rightsquigarrow \int_0^1 \left(\tilde{W}_d(r) - W(1)\int_r^1 (q-r)^{d-1}\mathrm{d}q\right)^2 \mathrm{d}r$$

  where $\tilde{W}_d(r) := \int_r^1 (q-r)^{d-1}\mathrm{d}W(q)$.

- We believe the result for part (i) should continue to hold for a much wider set of deterministic regressors, such as the conditions of Phillips and Xiao (1998) given on the previous slide. For part (iii) it should hold for the appropriate projection of $\mathbf{x}(r)$ under this type of conditions.
- For part (ii) the limiting null distribution of $(Tb_{d,T}^2)^{-1}S_{2,T}^{\mu}(d)$ again changes from that for the no deterministics case, and a formula is provided in the paper. It may be possible to obtain a representation for general $\mathbf{x}(r)$.

Deterministic Components

- Given $\mathbf{x}_t$ contains a constant, for the correlated components model considered earlier, the first derivative of the log-likehood can be shown to be constant with respect to the data and a LBUI test can be formed from the second derivative of the log likelihood.

- This LBUI test can be shown to coincide with the LBI test for the uncorrelated components model, $S_{T,\mathbf{x}}(d)$ in (14). Consequently in this case the approximate relation to Tanaka's fractional integration test statistic no longer holds.

Weak Dependence in ε_t

- The $S_T(d)$ statistic of (14) is LBI in the case where the observation error process $\{\varepsilon_t\} \sim NIID(0, \sigma_\varepsilon^2)$. Can it be modified to retain an asymptotically pivotal limiting null distribution, in cases where there is time series dependence in $\{\varepsilon_t\}$?
- In the $d = 1$ case Kwiatkowski *et al.* (1992) [KPSS] generalise (1) to the case where the observation error process $\{\varepsilon_t\}$ satisfies the α-mixing conditions of e.g. Phillips and Perron (1988, p.336), with long run variance

$$\sigma_L^2 := \lim_{T\to\infty} T^{-1} E\left(\sum_{t=1}^{T} \varepsilon_t\right)^2 .$$

- Under H_0 in the $d = 1$ case it holds that $(T\sigma_L)^{-2}\boldsymbol{y}'\boldsymbol{C}_1\boldsymbol{C}_1'\boldsymbol{y} \rightsquigarrow \int_0^1 W(r)^2 \mathrm{d}r$. To obtain an asymptotically pivotal statistic, KPSS suggest replacing σ_L^2 by the consistent estimator,

$$\hat{\sigma}_L^2 := \boldsymbol{y}'\boldsymbol{y}/T + 2\sum_{i=1}^{l} w(i,l)\boldsymbol{y}'\boldsymbol{y}_{-i}/T,$$

where $\boldsymbol{y}_{-i}$ is the vector $\boldsymbol{y}$ lagged i periods and $w(i,l) = 1 - i/(l+1)$, $i = 1, \ldots, l$.

Weak Dependence in ε_t

- As discussed in e.g. Stock (1994,p.2797), the bandwidth parameter l must be chosen to be $o(T^{1/2})$ to ensure that $\hat{\sigma}_L^2 \rightarrow^p \sigma_L^2$ under both H_0 and the local alternative H_c that the long run variance of η_t is $\sigma_L^2 c^2 / T^2$, where c is a non-negative finite constant.

- The KPPS-type statistic for $d = 1$ is therefore given by

$$S_T^L(1) := \frac{T^{-2} \boldsymbol{y}' \boldsymbol{C}_1 \boldsymbol{C}_1' \boldsymbol{y}}{\hat{\sigma}_L^2}$$

  whose limiting null distribution is $\int_0^1 W(r)^2 \mathrm{d}r$, exactly as in the case where ε_t is serially uncorrelated.

- For $d \in (1/2, \infty)$ the same approach can be validly used, so that

$$(T^2 b_{d,T}^2)^{-1} \frac{\boldsymbol{y}' \boldsymbol{C}_d \boldsymbol{C}_d' \boldsymbol{y}}{\hat{\sigma}_L^2} \rightsquigarrow \int_0^1 W_d^2(r) \mathrm{d}r.$$

Weak Dependence in ε_t

- For $d \in (1/4, 1/2]$ it is possible to correct the LBI statistic for serial correlation in ε_t but it is considerably more involved, so I will not discuss the details further.

- For $d \in (0, 1/4]$ things are considerably more problematic! This is also a feature of the LM fractional integration tests of Tanaka (1999) and Robinson (1994) where one has to essentially (correctly) specify a parametric model for the serial correlation in ε_t, eg $ARMA(p, q)$. This is currently work in (very slow!) progress.

- It is well know that the kernel-based serial correlation correction proposed in KPSS can be quite poor in finite samples, especially where ε_t is positively serially correlated, delivering considerably oversized tests. The finite sample efficacy of the non-parametric correction is expected to be poorer the smaller the value of d (essentially the smaller is d the more closely a positively autocorrelated weakly dependent process will resemble an $I(d)$ process).

Table of Contents

Numerical Simulations: Null Distributions

- "Simulated distributions":
  - We generated data under the null according to

$$y_t = \varepsilon_t \sim i.i.d.N(0,1), \quad t = 1, \ldots, T,$$

    and computed test statistics for $M = 10,000$ replications with $T = 1000$.
  - Based on these test statistics we computed kernel-density plots.
- "Limit distributions":
  - Normal distributions analytically.
  - Brownian functionals simulated.
- For now just mean-corrected statistics.
- We also tried other values of T with very similar-looking results.

Figure 1: Null distributions for $d \leq 0.5$

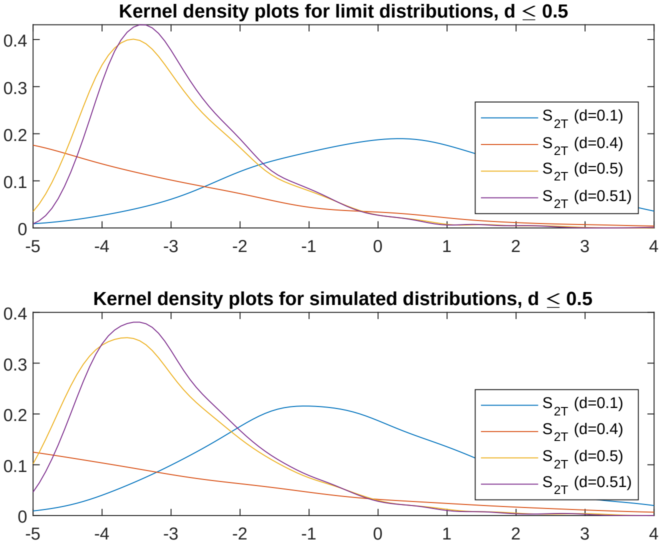
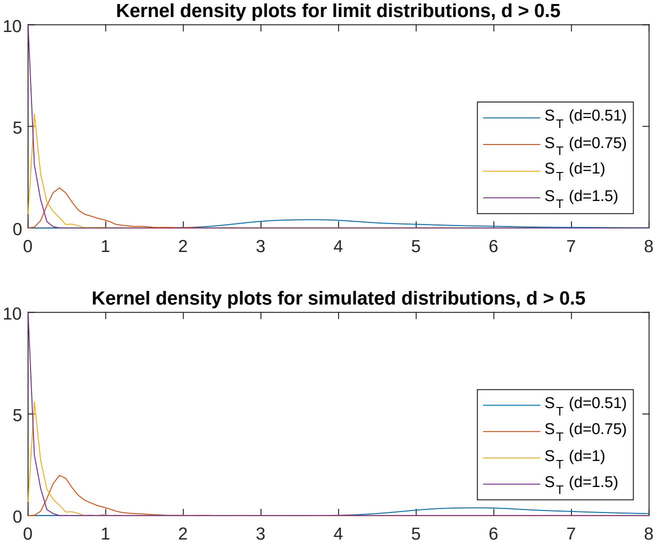


Figure 2: Null distributions for $d > 0.5$



Numerical Simulations: Power Functions

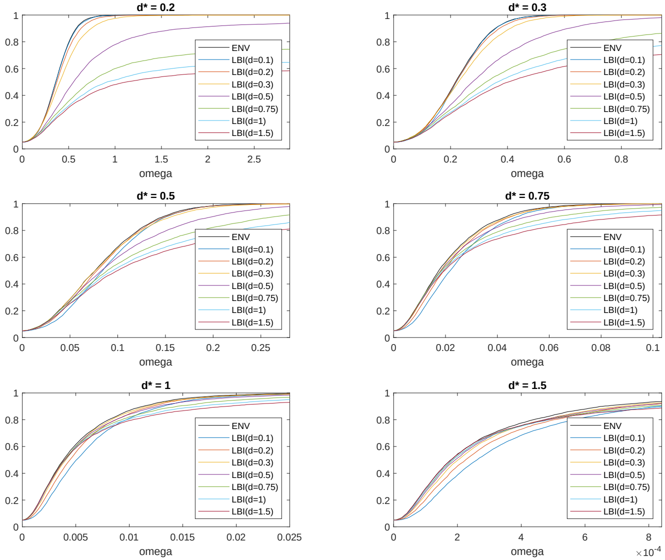
- We generated data under the alternative as

$$y_t = \Delta_+^{-d^*} \eta_t + \varepsilon_t, \quad t = 1, \ldots, T,$$
$$\epsilon \sim i.i.d.N(0,1), \quad \eta_t \sim i.i.d.N(0, \omega^2),$$

  and computed LBI(d) test statistics for $M = 10,000$ replications with $T = 1000$.
- For each pair (d^*, d) we calculated power as function of ω (for a range of relevant Pitman-type local alternatives).
- Note: d^* is a feature of the DGP, while d is chosen by the econometrician when calculating the LBI statistic.
- Both non-mean-corrected and mean-corrected statistics.
- We also tried other values of T with very similar-looking results.
- Also plotted (asymptotic local) power envelope for comparison.

Numerical Simulations: Power Functions

Figure 3: Power functions, non-mean-corrected statistics



Numerical Simulations: Power Functions

Figure 4: Power functions, mean-corrected statistics

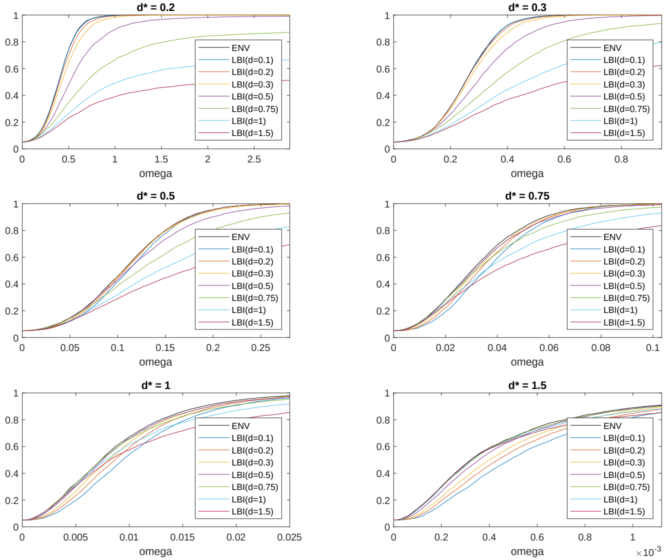


Table of Contents

Conclusions

- We develop a class of LBI tests for the null hypothesis that a time series is a martingale difference sequence against the alternative that it is a perturbed fractionally integrated (long memory) processes.
- The LBI tests are indexed by a user-chosen long memory parameter, $d > 0$, and are locally most powerful (under Gaussianity) where the true long memory parameter d^* coincides with this value.
- The class of tests contains a number of widely used tests as special cases, including the Nyblom and Mäkeläinen (1983) and Kwiatkowski *et al* (1992) [KPSS] tests.
- A taxonomy of limiting distribution theory for the class of statistics (indexed by d) is provided. These distributions depend on d, but can be straightforwardly simulated. In the paper we also explore the accuracy of the approximation provided by the asymptotic theory for finite T and propose a finite sample correction factor.
- Comparing the power properties of the LBI tests under appropriate Pitman drift sequences it is seen that they might better be called VLBI tests, V for very! Knowledge of d^* doesn't seem to help a huge amount.
- Extensions to allow for the presence of a general form of weak dependence under the null hypothesis and deterministic mean components are also considered.