European University Institute

Max Weber Programme

EUI Working Papers

MWP 2008/06

Fractional Integration and Cointegration Testing Using the Sample Mean

Matei Demetrescu

EUROPEAN UNIVERSITY INSTITUTE

MAX WEBER PROGRAMME

Fractional Integration and Cointegration Testing
Using the Sample Mean

MATEI DEMETRESCU

EUI Working Paper MWP No. 2008/06

ISSN 1830-7728

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy

http://www.eui.eu/
http://cadmus.eui.eu/

Abstract
The convergence rate of the sample mean of fractionally integrated processes is exploited to build test statistics for the fractional integration parameter *d* of univariate series, as well as for the rank of fractional cointegration of multivariate series with known or unknown order of fractional integration. Recursive adjustment is employed when dealing with deterministic components. The suggested test statistics are easy to compute and possess standard limiting distributions.

Keywords
Long memory testing, invariance principle, fractional Brownian motion

Fractional Integration and Cointegration Testing using the Sample Mean

Matei Demetrescu*
Goethe University Frankfurt†, European University Institute‡

1 Introduction

The most direct way to build a test for the fractional integration parameter d is to exploit the distribution of a suitable estimator. There is indeed a large body of literature concerning point estimation, in both the time and the frequency domain, such as Granger and Joyeux (1980) or Sowell (1992), and Geweke and Porter-Hudak (1983), Fox and Taqqu (1986) or Robinson (1995), respectively. Asymptotic distributions have been derived for these estimators (under several restrictions for d and additional regularity conditions), based on which tests may be built.

A different branch of the literature is concerned with the LM fractional integration test pioneered by Robinson (1991, 1994), which has asymptotic normal distribution. By working under the null hypothesis, the test does not require restrictions on d. Refinements have been proposed by Tanaka (1999), as well as Agiakloglou and Newbold (1994), Breitung and Hassler (2002) or, more recently, Demetrescu, Kuzin and Hassler (2008).

There are other possibilities to test hypotheses about d. In an autoregressive context, Dickey-Fuller type tests for the null of a unit root have

*E-mail addresses: deme@wiwi.uni-frankfurt.de, matei.demetrescu@eui.eu.
The author would like to thank Uwe Hassler, Vladimir Kuzin, Helmut Lütkepohl, Paulo M.M. Rodrigues and Adina I. Tarcolea for very helpful comments.
†Statistics and Econometric Methods, Goethe-University Frankfurt, Gräfstr. 78 PF 76, D-60054 Frankfurt, Germany.
‡Max Weber Programme, European University Institute, Via delle Fontanelle, 10, I-50014 San Domenico.

power against fractional alternatives with $d < 1$ under suitable assumptions (Diebold and Rudebusch, 1991, Hassler and Wolters, 1994, Krämer, 1998). A generalization following from this is the so-called fractional Dickey-Fuller test proposed by Dolado, Gonzalo and Mayoral (2002). Similarly, when testing the null of integration of order 0 [$I(0)$], the KPSS test due to Kwiatkowski *et al.* (1992) has power against $I(d)$ for $d \neq 0$ (Lee and Schmidt, 1996).

When cointegration testing is the focus, it may not be convenient to look at the long memory properties of residuals from a static regression, see the literature on residual log-periodogram inference, in particular Hassler, Marmol and Velasco (2006). While the multivariate LM test proposed by Nielsen (2005) or the (multivariate) test of Breitung and Hassler (2002) can be used to test the cointegration rank in the same manner as the Johansen (1995) trace test, they require known d, as does the residual-based cointegration test proposed by Hassler and Breitung (2006). Notably, Robinson and Hualde (2003) develop a procedure that works with estimated fractional integration parameters.

The testing approach presented here builds on the convergence rate of the sample mean of fractionally integrated processes, which depends on d. In fact, for stationary fractional integration, the suggested method parallels the use of a Central Limit Theorem [CLT] for the sample mean.

The contribution of this paper is threefold. First, extensions to test the rank of fractional cointegration are provided; it is shown that the test suggested here does not require knowledge about fractional integration order, since it copes with estimation of d under very weak requirements for the convergence rate of the used estimator $\widehat{d}$. Second, a way to deal with deterministic components is suggested. This is not a trivial issue, since one cannot apply a CLT to data from which the sample mean has been subtracted.[1] This non-trivial problem is avoided by using recursive demeaning (and detrending) when accounting for deterministics. Third, it is shown how to apply this approach to series exhibiting nonstationary fractional (co)integration.

The remainder of this paper is structured as follows. In Section 2, the suggested test is presented for stationary fractional integration, $d \in (-0.5, 0.5)$, followed by an examination of the cointegration case. Section 4 addresses the treatment of deterministic components. The extension to nonstationary

[1] The behavior of sample moments has been used before to test the null of $I(1)$, see the variance ratio test Phillips and Ouliaris (1990). Sample variances, however, can be computed from demeaned (or detrended) series.

fractional integration, $d \in (0.5, 1.5)$, is then dealt with. The final section concludes.

2 Stationary fractional integration

Let the observed series y_t be fractionally integrated of order d:

$$y_t = (1-L)^{-d} e_t, \quad t \in \mathbb{Z}, \tag{1}$$

where the fractional difference operator $(1-L)^d$ is given by the usual series expansion and $d \in (-0.5, 0.5)$. The short-memory process e_t is specified by the following assumption.

Assumption 1 *Let*

$$e_t = \varepsilon_t + \sum_{j \geq 1} b_j \varepsilon_{t-j},$$

where $\sum_{j \geq 1} |b_j| < \infty$, $\varepsilon_t \sim iid\,(0, \sigma^2)$ *and* $\exists s > 2$ *so that* $E\,|\varepsilon_t|^s < \infty$.

Let γ_h be the h^{th} autocovariance of e_t, and recall that the long-run variance of e_t, $\omega^2 = \gamma_0 + 2\sum_{h \geq 1} \gamma_h$, is finite and positive, $\omega^2 = \sigma^2 \left(1 + \sum_{j \geq 1} b_j\right)^2$.

An invariance principle holds under these assumptions (see e.g. McLeish, 1975):

$$\frac{1}{\omega T^{0.5+d}} \sum_{t=1}^{[sT]} y_t \Rightarrow B_d(s) \text{ as } T \to \infty, \tag{2}$$

where "$\Rightarrow$" stands for weak convergence in a suitable metric space, $[\cdot]$ is the floor function, and $B_d(s)$ a standard fractional Brownian motion of type I (as denoted by Marinucci and Robinson, 1999).[2]

One may define short-memory processes as processes for which weak convergence of their cumulated-sums process to a Brownian motion holds, see Lo (1991) for a short discussion. Similarly, in the context of fractional integration, one may extend this definition of short memory to the invariance principle in (2). The assumptions made on e_t are stronger than those required by McLeish (1975) and could be relaxed or replaced; in particular,

[2]Actually, a CLT would suffice for this section. An invariance principle, however, is needed later on, so using it from the beginning ensures a unified framework.

conditional heteroscedasticity could be allowed for. But discussing conditions under which an invariance principle as in (2) holds is not the aim of this paper and the general linear process assumption is kept as a simple way to illustrate the proposed tests.

The convergence rate of the sample mean depends on the fractional integration parameter d; it follows from (2) that

$$T^{0.5-d}\overline{y} \xrightarrow{d} \omega B_d(1),$$

where "$\xrightarrow{d}$" denotes convergence in distribution, and $\overline{y} = T^{-1}\sum_{t=1}^{T} y_t$. Multiplying the sample mean with $T^{0.5-d_0}$ hence leads under the null hypothesis $d = d_0$ to a proper asymptotic distribution. Under the alternative $d > d_0$, it obviously holds that $\left|T^{0.5-d_0}\overline{y}\right| \xrightarrow{p} \infty$; under the alternative $d < d_0$, $\left|T^{0.5-d_0}\overline{y}\right| \xrightarrow{p} 0$ holds, with "$\xrightarrow{p}$" standing for convergence in probability.

A pivotal test statistic for the null hypothesis $d = d_0$ is thus given by

$$\mathcal{T} = \frac{T^{1-2d_0}}{\widehat{\omega}^2}\overline{y}^2, \tag{3}$$

with $\widehat{\omega}^2$ a consistent estimator of the long-run variance of the short-memory component e_t. The following proposition then holds true:

Proposition 1 *Provided that $T \to \infty$, it holds for $\mathcal{T}$ from (3) with y_t from (1) and e_t from Assumption 1 under the null hypothesis $d = d_0$ that*

$$\mathcal{T} \xrightarrow{d} \chi_1^2.$$

Proof: Obvious and omitted.

One rejects in favor of $d < d_0$ for too small values, and in favor of $d > d_0$ for too large values of the test statistic, with the test obviously being consistent. One-sided as well as two-sided tests can be performed.

Remark 1 *Just as with the KPSS test, the divergence rate under the alternative hypothesis depends on the distance to the null hypothesis and, for alternatives very close to the null, it may be lower than the divergence rate of a test based on estimators of d. On the other hand, it may also be higher for large distances to the null. This behavior is the effect of focusing on such a general property of long-range dependent processes as the convergence rate of their sample mean.*

The common way to compute an estimate for the long-run variance is to use a kernel estimator of the spectral density of e_t at the origin (of the type studied by Newey and West, 1987, or Andrews, 1991) with the series differenced to stationarity.[3] Two possible estimation strategies arise:

First, one can estimate ω^2 under the null hypothesis, that is, based on $e_t = (1-L)^{d_0} y_t$. Under the alternative, $\widehat{\omega}^2$ is no longer consistent, and explodes if $d > d_0$ (implodes for $d < d_0$). This, however, does not affect the consistency of the test based on $\mathcal{T}$ from (3), since Lo (1991) and Lee and Schmidt (1996) show the relations $T^{2d_0-2d}\widehat{\omega}^2 \xrightarrow{p} 0$ for $d > d_0$ and $T^{2d_0-2d}\widehat{\omega}^2 \xrightarrow{p} \infty$ for $d < d_0$ to hold true for many kernel estimators $\widehat{\omega}$, leading to $\mathcal{T} \xrightarrow{p} 0$ for $d < d_0$ and $\mathcal{T} \xrightarrow{p} \infty$ for $d > d_0$.

Second, one can estimate ω^2 under the alternative hypothesis, that is, based on $\widehat{e}_t = (1-L)^{\widehat{d}} y_t$, with $\widehat{d}$ some consistent estimator for d. Since $\widehat{e}_t$ are in turn consistent estimators for e_t, ω^2 will itself be consistently estimated if the convergence rate of $\widehat{d}$ is higher than the rate at which the bandwidth increases, and the test retains the properties outlined in Proposition 1.

3 Stationary fractional cointegration

Let us now examine the case of multivariate time series. Assume that all K elements y_{tk} of the observed time series $\mathbf{y}_t$ have the same integration order d:

$$\mathbf{y}_t = (1-L)^{-d}\,\mathbf{e}_t,\ t \in \mathbb{Z} \tag{4}$$

with $\mathbf{e}_t \in \mathbb{R}^K$ a short memory process given by the following assumption:

Assumption 2 *Let*

$$\mathbf{e}_t = \boldsymbol{\varepsilon}_t + \sum_{j\geq 1} B_j \boldsymbol{\varepsilon}_{t-j},$$

where $\boldsymbol{\varepsilon}_t \sim iid\,(0,\Sigma)$ with Σ positive semidefinite, $\exists\, s > 2$ so that $E\,|\boldsymbol{\varepsilon}_t|^s < \infty$ and $\sum_{j\geq 1}\|B_j\| < \infty$, $\|\cdot\|$ being the matrix norm induced by the Euclidean vector norm.

Let Ω be the long-run covariance matrix of $\mathbf{e}_t$, $\Omega = B\Sigma B'$, with $B = I_K + \sum_{j\geq 1} B_j$, and recall that Ω is symmetric. The long-run covariance matrix

[3] There are also semiparametric approaches available, see e.g. Berk (1974), or subsampling-based ones, as introduced by Carlstein (1986).

Ω is positive definite if and only if y_{tk}, $k = 1, 2, \ldots, K$, are not cointegrated. Nielsen (2004) shows that Ω has exactly rank $K - r$, where r is the rank of fractional cointegration of $\mathbf{y}_t$.

In the case of no cointegration, the multivariate version of the test statistic from (3) is, naturally,

$$\mathcal{T}_K = T^{1-2d}\,\overline{\mathbf{y}}'\widehat{\Omega}^{-1}\overline{\mathbf{y}}.$$

For $r = r_0 > 0$, the probability limit (as $T \to \infty$) of $\widehat{\Omega}$ is singular and thus not invertible. This suggests using the Moore-Penrose inverse $\widehat{\Omega}^{-}$ instead of the usual inverse. Since the Moore-Penrose inverse is not continuous, special care needs to be taken in the estimation of Ω^{-}. That is, one needs to ensure that the rank of $\widehat{\Omega}^{-}$ converges a.s. to that of Ω^{-} (Andrews, 1987). This could be accomplished by estimating Ω by the usual methods and restricting its rank when building the inverse $\widehat{\Omega}^{-}$.

Furthermore, it turns out that the order of fractional integration need not be known: one may plug in an estimator $\widehat{d}$ without affecting the asymptotics. The test statistic is thus given by

$$\mathcal{T}_K = T^{1-2\widehat{d}}\,\overline{\mathbf{y}}'\widehat{\Omega}^{-}\overline{\mathbf{y}}, \tag{5}$$

and its behavior is characterized by the following proposition.

Proposition 2 *Under the null hypothesis $r = r_0$ for $\mathcal{T}_K$ from (5) with $\mathbf{y}_t$ from (4), $\mathbf{e}_t$ from Assumption 2, and $\widehat{d} = d + O_p\left(T^{-\alpha}\right)$ for some positive α, it holds as $T \to \infty$ that*

$$\mathcal{T}_K \xrightarrow{d} \chi^2_{K-r_0}.$$

Proof: See the Appendix.

When testing the null of $r = r_0$, there are two possible families of alternative hypotheses:

One can test the number of cointegration relations - the Johansen (1995) approach:

$$\begin{aligned} H_0 &: \; r = r_0 = 0, 1, \ldots, K-1 \\ H_1 &: \; r > r_0 \end{aligned}$$

Here, rejecting the null $r = r_0 = K - 1$ in favor of $r = K$ points toward a wrong fractional integration parameter d, i.e. the fractional integration parameter of (at least) one of the elements of $\mathbf{y}_t$ is smaller than d.

Following Nyblom and Harvey (2000), who analyze a multivariate version of the KPSS test,[4] one can alternatively test the number of common trends:

$$\begin{aligned} H_0 &: r = r_0 = K-1,\ldots,0 \\ H_1 &: r < r_0 \end{aligned}$$

Here, rejecting the null $r = r_0 = 0$ points toward a true fractional integration parameter that is larger than assumed for (at least) one of the univariate series y_{tk}.

The tests reject for too small values in the first family of tests, and for too large in the second.

Remark 2 *One uses the same test statistic in both types of tests and all null hypotheses $r = r_0$; only the limiting distributions change depending on r_0. Moreover, two-sided testing is straightforward to implement.*

Remark 3 *Rejection probabilities are easily computed for each null and alternative hypothesis. Thus, the properties of any estimation scheme of the cointegration rank based on sequential testing can be derived analytically, based on the cumulative distribution function of the χ^2 distribution(s).*

4 Accounting for deterministic components

Deterministic components, such as a non-zero mean, a linear time trend, and seasonally varying means, pose no problems when estimating ω^2 or d. This is not the case with the test statistics themselves, since usual demeaning would be an obvious mistake.

Let us discuss the non-zero mean case first, where a solution is offered by so-called recursive (adaptive) demeaning, introduced by So and Shin (1998). Assume one observes in the univariate case

$$y_t = z_t + \mu, \tag{6}$$

where z_t is fractionally integrated of order d with a short memory component as described by Assumption 1.

[4] Given the result of Lee and Schmidt (1996), the multivariate KPSS test should itself have power against fractional cointegration.

A handy solution is to apply the test to y_t^μ, the recursively demeaned y_t. The observations y_t from (6) are recursively demeaned as follows:

$$y_t^\mu = y_t - \frac{1}{t}\sum_{j=1}^{t} y_j,\ t = 2, 3, \ldots, T \tag{7}$$

and $y_1^\mu = 0$. A non-zero mean obviously cancels out, but the null distribution of the test statistic $\mathcal{T}$ computed with y_t^μ instead of y_t is changed. To be more precise, it requires a different standardization, since the following convergence result is easily established with the help of the Continuous Mapping Theorem under stationary fractional integration of order d, given the invariance principle in (2).

Lemma 1 *Define for some $c \in (0, 1)$*

$$B_d^\mu(s) = B_d(s) - \frac{1}{s}\int_0^s B_d(r)\, dr, \quad s \in [c, 1],$$

the recursively demeaned standard fractional Brownian motion. Then, provided that $T \to \infty$, it holds for y_t^μ from (7) that

$$\frac{1}{\omega T^{0.5+d}} \sum_{t=1}^{[sT]} y_t^\mu \Rightarrow B_d^\mu(s), \quad s \in [c, 1].$$

Proof: Obvious and omitted.

The recursively demeaned fractional Brownian motion is not defined for $s = 0$,[5] hence the restriction to a compact interval not containing 0. Fortunately, this is not a problem, since only what happens at $s = 1$ is of interest.

The recursively demeaned fractional Brownian motion has a different variance at $s = 1$, say $v_\mu(d) = Var(B_d^\mu(s))$. The test statistic then becomes

$$\mathcal{T}_\mu = \frac{T^{1-2d_0}}{v_\mu(d_0)\,\widehat{\omega}^2}\overline{y^\mu}^2. \tag{8}$$

[5] For $d = 0$, i.e. for the usual Brownian motion, the corresponding recursively demeaned Brownian motion possesses a proper limit as $s \to 0$, so one can extend it to a continuous process on $[0, 1]$ by setting $B_d^\mu(0) = 0$ almost surely (Chang, 2002, Section 5).

Since $B_d^\mu(s)$ is itself a Gaussian process, inheriting $B(s)$, it follows a normal distribution at $s = 1$. Hence, χ^2 asymptotics will follow for $\mathcal{T}_\mu$, a fact formalized in Proposition 3 below.

The standardizing factor $v_\mu(d)$ is not tractable analytically; numerical approximations are provided in the Appendix.

Should y_t contain a linear time trend,

$$y_t = z_t + \mu + \tau t, \tag{9}$$

the following recursive detrending scheme can be used

$$y_t^\tau = y_t + \frac{2}{t}\sum_{j=1}^{t} y_j - \frac{6}{t(t+1)}\sum_{j=1}^{t} j\, y_j,\ t = 2, 3, \ldots, T \tag{10}$$

and $y_1^\tau = 0$. Lemma 2 follows, similarly:

Lemma 2 *Define for some $c \in (0, 1)$*

$$B_d^\tau(s) = B_d(s) + \frac{2}{s}\int_0^s B_d(r)\, dr - \frac{6}{s^2}\int_0^s rB_d(r)\, dr, \quad s \in [c, 1],$$

the recursively detrended standard fractional Brownian motion continuous on $[0, 1]$. Then, provided that $T \to \infty$, it holds for y_t^τ from (10) that

$$\frac{1}{\omega T^{0.5+d}}\sum_{t=1}^{[sT]} y_t^\tau \Rightarrow B_d^\tau(s), \quad s \in [c, 1].$$

Proof: Obvious and omitted.

Again, this leads to a different standardizing factor, $v_\tau(d)$ for the test statistic

$$\mathcal{T}_\tau = \frac{T^{1-2d_0}}{v_\tau(d_0)\,\widehat{\omega}^2}\overline{y^\tau}^2. \tag{11}$$

Values of $v_\tau(d)$ are tabulated in the Appendix.

Moreover, Kuzin (2005) shows how to recursively deseasonalize an observed time series so that the problem is reduced to the case of recursive demeaning. The extension to deseasonalizing in the presence of a time trend follows directly from his work.

The proof of the following proposition is then easily established.

Proposition 3 *Provided that $T \to \infty$, it holds under the null hypothesis $d = d_0$ for $\mathcal{T}_\mu$ and $\mathcal{T}_\tau$ from (8), and (11), respectively, that*

$$\mathcal{T}_\mu \xrightarrow{d} \chi_1^2$$

and

$$\mathcal{T}_\tau \xrightarrow{d} \chi_1^2 .$$

Proof: Obvious and omitted.

Remark 4 *Consistency against the respective alternative hypotheses is guaranteed by the same mechanism as for the case without deterministic components.*

For multivariate time series, one can recursively remove the deterministic component for each element of $\mathbf{y}_t$ separately, since neither recursive demeaning or detrending affect the property of fractional cointegration. All elements, however, should be either recursively demeaned or recursively detrended:

$$\mathbf{y}_t^\mu = \mathbf{y}_t - \frac{1}{t}\sum_{j=1}^{t}\mathbf{y}_j, \ t = 2, 3, \ldots, T \tag{12}$$

$$\mathbf{y}_t^\tau = \mathbf{y}_t + \frac{2}{t}\sum_{j=1}^{t}\mathbf{y}_j - \frac{6}{t(t+1)}\sum_{j=1}^{t} j\,\mathbf{y}_j, \ t = 2, 3, \ldots, T, \tag{13}$$

with $\mathbf{y}_1^\mu = \mathbf{y}_1^\tau = 0$. The respective test statistics are

$$\mathcal{T}_{K\mu} = \frac{T^{1-2\widehat{d}}\,\overline{\mathbf{y}^\mu}'\widehat{\Omega}^{-}\overline{\mathbf{y}^\mu}}{v_\mu(d_0)}, \tag{14}$$

$$\mathcal{T}_{K\tau} = \frac{T^{1-2\widehat{d}}\,\overline{\mathbf{y}^\tau}'\widehat{\Omega}^{-}\overline{\mathbf{y}^\tau}}{v_\tau(d_0)}, \tag{15}$$

with $\widehat{d}$ and $\widehat{\Omega}$ computed with usual demeaning or detrending. The result analogous to Proposition 2 holds as well.

Proposition 4 *Provided that $T \to \infty$, it holds under the null hypothesis $r = r_0$ for $\mathcal{T}_{K\mu}$ and $\mathcal{T}_{K\tau}$ from (14), and (15), respectively, that*

$$\mathcal{T}_{K\mu} \xrightarrow{d} \chi_{K-r_0}^2$$

and

$$\mathcal{T}_{K\tau} \xrightarrow{d} \chi_{K-r_0}^2 .$$

Proof: Obvious and omitted.

5 Nonstationary fractional integration

In principle, the treatment of the nonstationary case does not pose additional difficulties. Rather, problems may arise from what one understands by nonstationary fractional integration since there are two ways of defining it. The starting point is the same short-memory component, but while one representation of an $I(d)$ process, $d > 0.5$, is that of a cumulated-sums process where the increments are stationarily integrated of order $d-1$, the other one is given by

$$y_t = (1-L)^{-d} e_t, \quad t \in \{1, 2, \ldots\},$$

with $y_t = 0$ for $t \leq 0$.[6] The test idea nevertheless translates in the same manner for both definitions.

Begin by assuming the cumulated-sums process to be the true data generating process. The second case is briefly addressed at the end of the section.

In the univariate case, it holds because of (2) that

$$\frac{1}{\omega T^{d-0.5}} y_{[sT]} \Rightarrow B_{d-1}(s). \tag{16}$$

Under fractional integration of order d, it follows with the Continuous Mapping Theorem that

$$\frac{1}{\omega T^{d-0.5}} \overline{y} \xrightarrow{d} \int_0^1 B_{d-1}(s)\, ds.$$

The integral follows a normal distribution; the variance of the integral, however, does not equal unity. Thus, a standardizing factor $v(d)$ in the case of nonstationary fractional integration is necessary even without having recursively demeaned/detrended the data. For $d \in (0.5, 1.5)$, the variance function $v(d) = Var\left(\int_0^1 B_{d-1}(s) ds\right)$ is given by

$$\begin{aligned} v(d) &= \int_0^1 \int_0^1 Cov\left(B_{d-1}(s), B_{d-1}(v)\right) dsdv \\ &= \frac{1}{2} \int_0^1 \int_0^1 \left(s^{2d-1} + v^{2d-1} - |s-v|^{2d-1}\right) dsdv \\ &= \frac{1}{1+2d}. \end{aligned}$$

[6] Marinucci and Robinson (1999) note that the first definition tends to be used in the probabilistic literature and the second one is preferred by econometricians.

For consistency of notation, one can define $v(d) = 1$ for $d \in (-0.5, 0.5)$.[7] Then, for stationary as well as nonstationary fractional integration, one can write

$$\mathcal{T} = \frac{T^{1-2d_0}}{v\left(d_0\right)\widehat{\omega}^2}\overline{y}^2,$$

with

$$\mathcal{T} \sim \chi_1^2.$$

Note that a nonzero mean is automatically accommodated, due to division by T^{2d_0-1}, since $d_0 > 0.5$. The same holds for structural breaks or seasonally varying means. But for large values of the mean, size or power distortions will appear in small samples. Moreover, a linear trend does need to be removed, in contrast to a non-zero mean.

A way to eliminate these is again offered by recursive removal of deterministics. When computing $\mathcal{T}_\mu$ and $\mathcal{T}_\tau$ based on recursively demeaned observations, Lemma 3 follows:

Lemma 3 *Under Assumption 1, it holds with y_t^μ from (7) and y_t^τ from (10) and $d \in (0.5, 1.5)$ as $T \to \infty$ that*

$$\frac{1}{\omega T^{d-0.5}}\overline{y^\mu} \xrightarrow{d} \int_0^1 \left(1 + \ln s\right) B_{d-1}\left(s\right) ds,$$

$$\frac{1}{\omega T^{d-0.5}}\overline{y^\tau} \xrightarrow{d} \int_0^1 \left(6s - 2\ln s - 5\right) B_{d-1}\left(s\right) ds.$$

Proof: See the Appendix.

Remark 5 *It is true that such representations could be derived for the stationary integration case as well; they would, however, lead to Stjeltjes-type integrals over $B_d(s)$. For $d = 0$ this poses no problems, but otherwise it is unclear how to define such integrals for $d \neq 0$, see the discussion in Pipiras and Taqqu (2003).*

The standardizing factors become

$$v_\mu(d) = \int_0^1 \int_0^1 \left(1 + \ln s\right)\left(1 + \ln v\right) Cov\left(B_{d-1}(s), B_{d-1}(v)\right) dsdv$$

[7] Note the discontinuity at $d = 0.5$.

and

$$v_\tau(d) = \int_0^1 \int_0^1 (6s - 2\ln s - 5)(6v - 2\ln v - 5)\, Cov\,(B_{d-1}(s), B_{d-1}(v))\, dsdv,$$

where, as before,

$$Cov\,(B_{d-1}(s), B_{d-1}(v)) = \frac{1}{2}\left(s^{2d-1} + v^{2d-1} - |s-v|^{2d-1}\right).$$

In this case, integration can be analytically dealt with, but the result is expressed with hyperbolic functions, which are not user-friendly. Again, numerical approximations are provided in the Appendix.

The extension to the multivariate case is straightforward, leading to the same test statistic as in (5), (14) and (15), respectively, with the same asymptotic properties.

Should one opt for the alternative definition of nonstationary fractional integration, the employed invariance principles still hold, but in terms of standard fractional Brownian motions of type II. Marinucci and Robinson (2000) show such convergence to hold under an additional assumption on the short-memory component, $\sum_{k\geq 0}\sum_{j\geq k+1}\left(\|B_j\|^2 + \|B_{-j}\|^2\right) < \infty$.

For the perhaps more empirically relevant case $d = 1$, $W_{d-1}(s)$ and $B_{d-1}(s)$ coincide with the standard Wiener process. With respect to the suggested test, the main difference between $W_{d-1}(s)$ and $B_{d-1}(s)$ for the case $d \neq 1$ lies in the respective covariance functions, where the fractional Brownian motion of type II exhibits more dependence in the neighborhood of the origin. The difference however is only significant for infinitesimally small distances (see Marinucci and Robinson, 1999), so using the standardizing factors v, v_μ and v_τ is justifiable. The test statistics from (5), (14) and (15) remain otherwise unaffected.

6 Summary

A family of tests for fractional integration and cointegration has been proposed, tests that are based on the convergence rate of the sample mean.

The tests are easy to compute and possess χ^2 limiting distributions. In the case of cointegration testing, knowledge of the fractional integration parameter is not required.

Recursive schemes for the removal of deterministic components are used, with the test statistics retaining their standard asymptotics.

Appendix

A Tabulated values of $v(d)$, $v_\mu(d)$, and $v_\tau(d)$

The numerical approximations provided for the stationary case are based on the following expression for $Var\left(B_d^\mu(1)\right) = v_\mu(d)$:

$$\lim_{T\to\infty} \frac{1}{T^{1+2d}} \sum_{t=2}^{T}\sum_{u=2}^{T} \frac{\Gamma(1-2d)\Gamma(|t-u|+d)}{\Gamma(d)\Gamma(1-d)\Gamma(1+|t-u|-d)} \times \left(1+\ln\frac{t-1}{T}\right)\left(1+\ln\frac{u-1}{T}\right).$$

This is a consequence of

$$\frac{1}{T^{0.5+d}}\sum_{t=1}^{T} y_t^\mu = \frac{1}{T^{0.5+d}}\sum_{t=2}^{T} y_t\left(1+\ln\frac{t-1}{T}\right) + o_p(1),$$

which is established with arguments similar to those in the proof of Lemma 3. The required variance of the recursively demeaned fractional Brownian motion at $s=1$ is then

$$Var(B_d^\mu(1)) = \lim_{T\to\infty} Var\left(\frac{1}{T^{0.5+d}}\sum_{t=2}^{T} y_t\left(1+\ln\frac{t-1}{T}\right)\right).$$

This can be expressed as a function of variances and covariances of y_t. At this point, one may assume a fractional white noise model for y_t, since the corresponding invariance principle guarantees the expression above to have the same limit, irrespective of the form of the short memory component. It follows that

$$Var(B_d^\mu(1)) = \lim_{T\to\infty}\frac{1}{T^{1+2d}}\left(\sum_{t=2}^{T}\sum_{u=2}^{T}\gamma_d(|t-u|)\left(1+\ln\frac{t-1}{T}\right)\left(1+\ln\frac{u-1}{T}\right)\right),$$

with the h^{th} autocovariance of the fractionally integrated standard white noise,

$$\gamma_d(h) = \frac{\Gamma(1-2d)\Gamma(h+d)}{\Gamma(d)\Gamma(1-d)\Gamma(1+h-d)}.$$

For recursive detrending, the following expression for $Var\left(B_d^{\tau}(1)\right)$ is obtained with the same arguments as above:

$$\lim_{T\to\infty} \frac{1}{T^{1+2d}} \sum_{t=2}^{T}\sum_{u=2}^{T} \frac{\Gamma(1-2d)\Gamma(|t-u|+d)}{\Gamma(d)\Gamma(1-d)\Gamma(1+|t-u|-d)}$$
$$\times\left(-5-2\ln\frac{t-1}{T}+6\frac{t}{T+1}\right)\left(-5-2\ln\frac{u-1}{T}+6\frac{u}{T+1}\right).$$

For computing the approximation, T is set equal to 12 000. This ensures a two-digit precision for the results.[8]

For the nonstationary case, numerical integration is used to compute $v_{\mu}(d)$ and $v_{\tau}(d)$. Computation is carried out with Euler's method and a number of $N = 1000$ evaluation points.

The results are given in Table 1 (together with $v(d)$ for completeness).

B Proofs

Proof of Proposition 2

The invariance principle given in (2) has a multivariate extension, proven by Csörgö and Mielniczuk (1995):[9]

$$\frac{1}{T^{d-0.5}} \sum_{t=1}^{[sT]} \mathbf{y}_t \Rightarrow \Omega^{0.5}\mathbf{B}_d(s), \tag{17}$$

where $\mathbf{B}_d(s)$ is a K-dimensional vector of independent standard fractional Brownian motions of type I and "$\Rightarrow$" stands for joint weak convergence.

[8] The precision can be assessed by examining the case $d = 0$, where it can be shown analytically that $v_{\mu}(0) = v_{\tau}(0) = 1$.

[9] Again, they pose less restrictive conditions than the general linear process assumed here.

Table 1: Standardizing factors $v\,(d)$, $v_\mu\,(d)$, and $v_\tau\,(d)$

d	$v\,(d)$	$v_\mu\,(d)$	$v_\tau\,(d)$	d	$v\,(d)$	$10^2 v_\mu\,(d)$	$10^2 v_\tau\,(d)$
-0.48	1.00	85.67	290.75	0.52	0.49	1.10	0.48
-0.46	1.00	63.72	207.75	0.54	0.48	2.02	0.84
-0.44	1.00	47.80	149.34	0.56	0.47	2.81	1.12
-0.42	1.00	36.18	108.06	0.58	0.46	3.50	1.35
-0.40	1.00	27.66	78.75	0.60	0.45	4.10	1.53
-0.38	1.00	21.37	57.84	0.62	0.45	4.62	1.66
-0.36	1.00	16.68	42.84	0.64	0.44	5.06	1.77
-0.34	1.00	13.16	32.01	0.66	0.43	5.45	1.84
-0.32	1.00	10.50	24.15	0.68	0.42	5.79	1.89
-0.30	1.00	8.48	18.40	0.70	0.42	6.08	1.92
-0.28	1.00	6.91	14.17	0.72	0.41	6.33	1.93
-0.26	1.00	5.70	11.02	0.74	0.40	6.55	1.92
-0.24	1.00	4.75	8.66	0.76	0.40	6.73	1.91
-0.22	1.00	4.00	6.88	0.78	0.39	6.88	1.88
-0.20	1.00	3.40	5.53	0.80	0.38	7.01	1.85
-0.18	1.00	2.92	4.48	0.82	0.38	7.12	1.81
-0.16	1.00	2.52	3.67	0.84	0.37	7.21	1.76
-0.14	1.00	2.20	3.04	0.86	0.37	7.28	1.71
-0.12	1.00	1.93	2.53	0.88	0.36	7.33	1.66
-0.10	1.00	1.70	2.13	0.90	0.36	7.38	1.60
-0.08	1.00	1.52	1.80	0.92	0.35	7.41	1.54
-0.06	1.00	1.35	1.54	0.94	0.35	7.43	1.48
-0.04	1.00	1.22	1.32	0.96	0.34	7.44	1.42
-0.02	1.00	1.10	1.14	0.98	0.34	7.44	1.36
0.00	1.00	1.00	1.00	1.00	0.33	7.43	1.30
0.02	1.00	0.91	0.87	1.02	0.33	7.42	1.23
0.04	1.00	0.83	0.76	1.04	0.32	7.40	1.17
0.06	1.00	0.76	0.67	1.06	0.32	7.38	1.11
0.08	1.00	0.70	0.60	1.08	0.32	7.35	1.05
0.10	1.00	0.64	0.53	1.10	0.31	7.32	0.99
0.12	1.00	0.59	0.47	1.12	0.31	7.28	0.93
0.14	1.00	0.55	0.42	1.14	0.30	7.24	0.87
0.16	1.00	0.51	0.38	1.16	0.30	7.20	0.81
0.18	1.00	0.47	0.34	1.18	0.30	7.15	0.75
0.20	1.00	0.44	0.31	1.20	0.29	7.10	0.69
0.22	1.00	0.41	0.28	1.22	0.29	7.06	0.64
0.24	1.00	0.39	0.25	1.24	0.29	7.01	0.59
0.26	1.00	0.36	0.23	1.26	0.28	6.95	0.53
0.28	1.00	0.34	0.21	1.28	0.28	6.90	0.48
0.30	1.00	0.32	0.19	1.30	0.28	6.85	0.43
0.32	1.00	0.30	0.17	1.32	0.27	6.79	0.38
0.34	1.00	0.28	0.16	1.34	0.27	6.73	0.34
0.36	1.00	0.27	0.14	1.36	0.27	6.68	0.29
0.38	1.00	0.25	0.13	1.38	0.27	6.62	0.24
0.40	1.00	0.24	0.12	1.40	0.26	6.56	0.20
0.42	1.00	0.22	0.11	1.42	0.26	6.51	0.16
0.44	1.00	0.21	0.10	1.44	0.26	6.45	0.12
0.46	1.00	0.20	0.09	1.46	0.26	6.39	0.08
0.48	1.00	0.19	0.09	1.48	0.25	6.33	0.04

Note. For $d > 0.5$, values are given for $100\,v_\mu\,(d)$ and $100\,v_\tau\,(d)$. See the text for details on how the figures were obtained.

It is known that, if $X \sim N(0, \Sigma)$, $X'\Sigma^{-}X \sim \chi^2_{\text{rank}(\Sigma)}$. $\mathbf{B}_d(1)$ being a vector of standard normal independent variables, the result follows directly for known d when ensuring consistent estimation of Σ^{-}.

Denote $\mathcal{T}_K^*$ the test statistic computed with known d. When the fractional integration order is estimated, $\mathcal{T}_K = \mathcal{T}_K^* \cdot T^{2d-2\widehat{d}}$ holds. Since $\lim_{T\to\infty} T^{-2T^{-\alpha}} = \lim_{T\to\infty} \exp(-2T^{-\alpha}\ln T) = 1$, the desired result follows.

Proof of Lemma 3

It holds for $d \in (0.5, 1.5)$ that

$$T^{0.5-d}\overline{y^{\mu}} = \frac{1}{T^{0.5+d}}\sum_{t=2}^{T}\left(y_t - \frac{1}{t}\sum_{j=1}^{t} y_j\right) = \frac{1}{T^{0.5+d}}\sum_{t=2}^{T} y_t\left(1 - \sum_{j=t}^{T}\frac{1}{j}\right).$$

Since $\sum_{j=1}^{p}\frac{1}{j} = C + \ln p + O\left(\frac{1}{p}\right)$, with C Euler's constant, one can write for $t \geq 2$

$$\begin{aligned}\sum_{j=t}^{T}\frac{1}{j} &= \sum_{j=1}^{T}\frac{1}{j} - \sum_{j=1}^{t-1}\frac{1}{j} = \ln T + O\left(\frac{1}{T}\right) - \ln(t-1) - O\left(\frac{1}{t-1}\right) \\ &= \ln\left(\frac{T}{t-1}\right) + O\left(\frac{1}{T} - \frac{1}{t-1}\right).\end{aligned}$$

It follows that

$$\frac{1}{T^{0.5+d}}\sum_{t=2}^{T} y_t^{\mu} = \frac{1}{T^{0.5+d}}\sum_{t=2}^{T} y_t\left(1 + \ln\frac{t-1}{T}\right) + O_p\left(\frac{1}{T^{0.5+d}}\sum_{t=2}^{T}\frac{y_t}{t-1}\right).$$

For the first term, it follows from the Continuous Mapping Theorem that

$$\frac{1}{T^{0.5+d}}\sum_{t=2}^{T} y_t\left(1 + \ln\frac{t-1}{T}\right) \Rightarrow \omega\int_0^1 (1 + \ln s)\, B_{d-1}(s)\, ds.$$

For the second term, it holds due to (16) that $y_t = O_p(T^{d-0.5})$, or $T^{0.5-d}y_t = O_p(1)$. Then,

$$\frac{1}{T^{0.5+d}}\sum_{t=2}^{T}\frac{y_t}{t-1} = \frac{1}{T}\sum_{t=2}^{T}\frac{T^{0.5-d}y_t}{t-1} = O_p\left(\frac{1}{T}\sum_{t=2}^{T}\frac{1}{t-1}\right) = O_p\left(\frac{\ln T}{T}\right),$$

so this term converges to zero for positive d. For $d \leq 0$, one proceeds along the same lines to obtain the first result.

When recursively detrending, one has

$$T^{0.5-d}\overline{y^{\tau}} = \frac{1}{T^{0.5+d}} \sum_{t=2}^{T} y_t^{\tau} = \frac{1}{T^{0.5+d}} \sum_{t=2}^{T} \left(y_t + \frac{2}{t} \sum_{j=1}^{t} y_j - \frac{6}{t(t+1)} \sum_{j=1}^{t} j\, y_j \right).$$

Rearranging terms leads to

$$T^{0.5-d}\overline{y}^{\tau} = \frac{1}{T^{0.5+d}} \sum_{t=2}^{T} y_t \left(1 + 2 \sum_{j=t}^{T} \frac{1}{j} - 6t \sum_{j=t}^{T} \frac{1}{j(j+1)} \right).$$

Since

$$\sum_{j=t}^{T} \frac{1}{j(j+1)} = \sum_{j=t}^{T} \left(\frac{1}{j} - \frac{1}{j+1} \right) = \frac{1}{t} - \frac{1}{T+1},$$

it follows as before that

$$\frac{1}{T^{0.5+d}} \sum_{t=2}^{T} y_t^{\tau} = \frac{1}{T^{0.5+d}} \sum_{t=2}^{T} y_t \left(1 - 2 \ln \frac{t-1}{T} - 6 + 6 \frac{t}{T+1} \right) + o_p(1),$$

from which the second result follows.

References

Agiakloglou, C., Newbold, P. (1994), Lagrange Multiplier Tests for Fractional Difference; *Journal of Time Series Analysis* 15, 253-262.

Andrews, D.W.K. (1987), Asymptotic Results for Generalized Wald Tests; *Econometric Theory* 3, 348-358.

Andrews, D.W.K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation; *Econometrica* 59, 817-858.

Berk, K.N. (1974), Consistent Autoregressive Spectral Estimates; *The Annals of Statistics* 2, 489-502.

Breitung, J., Hassler, U. (2002), Inference on the Cointegration Rank in Fractionally Integrated Processes; *Journal of Econometrics* 110, 167-185.

Carlstein, E. (1986), The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence; *The Annals of Statistics* 14, 1171-1179.

Chang, Y. (2002), Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency; *Journal of Econometrics* 110, 261-292.

Csörgö, S., Mielniczuk, J. (1995), Distant Long-range Dependent Sums and Regression Estimation; *Stochastic Processes and their Applications* 59, 143-155.

Demetrescu, M., Kuzin, V., Hassler, U. (2008), Long Memory Testing in the Time Domain; *Econometric Theory* 24, 176-215.

Diebold, F.X., Rudebusch, G.D. (1991), On the Power of Dickey-Fuller Tests against Fractional Alternatives; *Economics Letters* 35, 155-160.

Dolado, J.J., Gonzalo, J., Mayoral, L. (2002), A Fractional Dickey-Fuller Test for Unit Roots; *Econometrica* 70, 1963-2006.

Fox, R., Taqqu, M.S. (1986), Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series; *Annals of Statistics* 14, 517-532.

Geweke, J., Porter-Hudak, S. (1983), The Estimation and Application of Long Memory Time Series; *Journal of Time Series Analysis* 4, 221-238.

Granger, C.W.J., Joyeux, R. (1980), An Introduction to Long Memory Time Series Models and Fractional Differencing; *Journal of Time Series Analysis* 1, 15-39.

Hassler, U., Breitung, J. (2006), A Residual-based LM Type Test against Fractional Cointegration; *Econometric Theory* 22, 1091-1111.

Hassler, U., Marmol, F., Velasco, C. (2006), Residual Log-Periodogram Inference for Long-run Relationships; *Journal of Econometrics* 130, 165-207.

Hassler, U., Wolters, J. (1994), On the Power of Unit Root Tests against Fractionally Integrated Alternatives; *Economics Letters* 45, 1-5.

Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*; Oxford University Press, New York.

Krämer, W. (1998), Fractional Integration and the Augmented Dickey-Fuller Test; *Economics Letters* 61, 269-272.

Kuzin, V. (2005), Recursive Demeaning and Deterministic Seasonality; *Statistics & Probability Letters* 72, 195-204.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. , Shin, Y. (1992), Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root; *Journal of Econometrics* 54, 159-178.

Lee, D., Schmidt, P. (1996), On the Power of the KPSS Test of Stationarity against Fractionally-integrated Alternatives; *Journal of Econometrics* 73, 285-302.

Lo, A.W. (1991), Long-Term Memory in Stock Market Prices; *Econometrica* 59, 1279-1313.

Marinucci, D., Robinson, P.M. (1999), Alternative Forms of Fractional Brownian Motion; *Journal of Statistical Planning and Inference* 80, 111-122.

Marinucci, D., Robinson, P.M. (2000), Weak Convergence of Multivariate Fractional Processes; *Stochastic Processes and their Applications* 86, 103-120.

McLeish, D.L. (1975), Invariance Principles for Dependent Variables; *Probability Theory and Related Fields* 32, 165 - 178.

Newey, W.K., West, K.D. (1987), A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix; *Econometrica* 55, 703-708.

Nielsen, M.O. (2004), Spectral Analysis of Fractionally Cointegrated Systems; *Economics Letters* 83, 225-231.

Nielsen, M.O. (2005), Multivariate Lagrange Multiplier Tests for Fractional Integration; *Journal of Financial Econometrics* 3, 372-398.

Nyblom, J., Harvey, A. (2000), Tests of Common Stochastic Trends; *Econometric Theory* 16, 176-199.

Phillips, P.C.B., Ouliaris, S. (1990) Asymptotic Properties of Residual Based Tests for Cointegration; *Econometrica* 58, 165-193.

Pipiras, V., Taqqu, M.S. (2003) Fractional Calculus and its Connections to Fractional Brownian Motion. In Doukhan, P., Oppenheim, G., Taqqu, M.S. (eds.), *Long-range Dependence: Theory and Applications*; Birkhäuser, Boston, 165-201.

Robinson, P.M. (1991), Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regression; *Journal of Econometrics* 47, 67-84.

Robinson, P.M. (1994), Efficient Tests of Nonstationary Hypotheses; *Journal of the American Statistical Association* 89, 1420-1437.

Robinson, P.M. (1995), Gaussian Semiparametric Estimation of Long Range Dependence; *Annals of Statistics* 23, 1630-1661.

Robinson, P.M., Hualde, J. (2003), Cointegration in Fractional Systems with Unknown Integration Orders; *Econometrica* 71, 1727-1766.

So, B.S., Shin, D.W. (1998), Recursive Mean Adjustment in Time-series Inferences; *Statistics & Probability Letters* 43, 65-73.

Sowell, F. (1992), Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models; *Journal of Econometrics* 53, 165-188.

Tanaka, K. (1999), The Nonstationary Fractional Unit Root; *Econometric Theory* 15, 549-582.