ELSEVIER

Journal of Econometrics 105 (2001) 225–247

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Semiparametric fractional cointegration analysis

D. Marinucci[a], P.M. Robinson[b,*]

[a] Dipartimento di Studi Geoeconomici, Statistici, Storici, University La Sapienza, Via del Castro Laurenziano 9, Quarto Piano, 00161 Roma, Italy

[b] Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK

Abstract

Fractional cointegration is viewed from a semiparametric viewpoint as a narrow-band phenomenon at frequency zero. We study a narrow-band frequency domain least squares estimate of the cointegrating vector, and related semiparametric methods of inference for testing the memory of observables and the presence of fractional cointegration. These procedures are employed in analysing empirical macroeconomic series; their usefulness and feasibility in finite samples is supported by results of a Monte Carlo experiment.

JEL classification: C22

Keywords: Semiparametric analysis; Fractional cointegration

1. Introduction

Modelling of long term relationships between macroeconomic variables has mostly centred around the possibility of cointegration of time series with one or more autoregressive (AR) unit roots. The unobservable input to the finite-degree vector AR is typically a vector $I(0)$ process, namely one that is (covariance) stationary with spectral density matrix that is continuous and positive definite at frequency zero. If single differencing of the AR observable, denoted by the $p \times 1$ column vector z_t, $t \in Z$, $Z = \{t\colon t = 0, \pm 1, \ldots\}$ produces

* Corresponding author. Tel.: +44-0171-955-7509; fax: +44-0171-831-1840.
E-mail address: p.m.robinson@lse.ac.uk (P.M. Robinson).

0304-4076/01/$ - see front matter
PII: S0304-4076(01)00076-8

an $I(0)$ process, z_t is said to be in $I(1)$, if twice differencing, $I(2)$, and so on; the argument of I is referred to as the integration order. The input process may have parametric autocorrelation, in an effort to simultaneously model short-run behaviour, or it may have nonparametric autocorrelation, recognizing that misspecification of even short-run behaviour can invalidate inferences on long-run behaviour; the modelling of z_t is then said to be, respectively, parametric or semiparametric. Empirical analysis typically begins with testing for the presence of unit roots; given a positive outcome, there is a search for possible cointegrating relationships, namely linear combinations of form

$$\alpha' z_t = e_t, \tag{1}$$

where α' is the transpose of the $p \times 1$ vector α, usually unknown, and e_t has a lower integration order than z_t. Such e_t, referred to as a cointegrating error, is $I(0)$ when z_t is $I(1)$, and is $I(0)$ or $I(1)$ when z_t is $I(2)$. Work on the parametric $I(1)$ observable case began with Engle and Granger (1987) and Johansen (1991); for semiparametric models, see e.g. Phillips (1991a, b).

The AR-based unit root testing and cointegration methods have been widely applied, suggesting that many economic time series could be $I(1)$ or $I(2)$, and providing information on the presence or absence of cointegration in many data sets. However, $I(1)$ and $I(2)$ are specialized forms of nonstationarity, while $I(0)$ is a specialized form of stationarity. In particular, scalar $I(0)$ processes are also nested within a much more general stationary and invertible fractional $I(d)$ class, for $|d| < \frac{1}{2}$, defined in the following section, such that the spectral density behaves like λ^{-2d} near $\lambda = 0$, λ denoting frequency, so the spectrum has a pole when $0 < d < \frac{1}{2}$, or a zero when $-\frac{1}{2} < d < 0$, at $\lambda = 0$. We call d the integration order; $I(d)$ nonstationary sequences, for $d \geqslant \frac{1}{2}$, can be defined such that their integer difference of suitable degree is a stationary and invertible (for $d > \frac{1}{2}$) fractional, or as suitable filter of an $I(0)$ sequence, as described in the following section. It is then possible that, for example, a test that a macroeconomic variable is $I(1)$ directed against fractional $I(d)$ alternatives might produce a different outcome from one directed against the usual stationary AR alternatives. Further, fractional processes might better approximate either z_t or e_t or both, and cointegration of stationary $I(d)$ processes can be entertained, where e_t is stationary $I(d_e)$ for $d_e < d$, and may be of interest in some financial series exhibiting long range dependence.

Now that large sample rules of inference for fractional $I(d)$ processes are available, analysing fractional integration and cointegration is a realistic possibility. For parametric stationary $I(d)$ processes, asymptotic theory of Fox and Taqqu (1986) has been extended by various subsequent authors, while for nonparametric stationary $I(d)$ processes (where the spectral density is unrestricted away from zero frequency) Robinson (1995a, b) has established asymptotic distributional properties of estimates of d, his results extended to nonstationary series by Velasco (1999a, b); limit distributions are standard. Notions of fractional cointegration have been explored, indeed the early paper of Engle and Granger (1987) stressing AR-based cointegration included a definition covering fractional processes; the stationary case has been studied by Robinson (1994a), the nonstationary one by Chan and Terrin (1995) and others.

Fractional modelling considerably expands the possibilities of cointegration analysis and poses considerable new challenges. The various methods developed for AR-based cointegration analysis depend on the presumed, integer-valued, integration orders of z_t and e_t, and appear to lose validity when the true integration orders differ. Such methods may be generalizable to pre-specified alternative, possibly fractional, integration orders, but faced with an uncountable infinity of possible integration orders it may be hard to choose ones even to be the subject of a pre-test. Generally, it seems more natural to allow integration orders to be unknown. This constitutes a radical departure from the AR-based approach, where integration orders, after testing, are treated as given. Additional complications that arise in the fractional setting are the possibility of a variety of integration orders in the vector z_t and, when it also is a vector, the cointegrating error e_t. Study of identification problems, of testing for the presence and degree of cointegration, and inference on the unknown coefficients of the cointegrating errors, is in its infancy.

The present paper develops and numerically evaluates methodology for inference on (possibly fractional) integration orders and (possibly fractional) cointegration, and for estimation of the cointegrating vector. In the following section, we discuss notions of fractional integration and cointegration. The cointegrating regression vector estimates are those of Robinson and Marinucci (1999a, b), denoted narrow-band frequency domain least squares (FDLS) and described in Section 3. Section 4 discusses estimates and test procedures in relation to integration orders of Robinson (1995a, b), Lobato (1996), Lobato and Robinson (1998), and proposes a related, Hausman-type test for the presence of fractional cointegration. Section 5 applies the procedures of Sections 3 and 4 to macroeconomic data sets used in the early papers of Engle and Granger (1987) and Campbell and Shiller (1987). The emphasis here is on gaining information on whether cointegration exists, which requires in the first place testing whether observable series have equal integration orders. In Section 6, Monte Carlo simulations assess finite sample performance of our Hausman-type test and compare various estimates of cointegrating coefficients, a main question there being how the differing theoretical results on asymptotic (higher order) bias of various estimates are relevant in finite samples. In view of our theoretical discussion and the empirical results, we recommend semiparametric frequency domain statistics in general, and the FDLS and Hausman-type test in particular, as feasible and useful procedures for general fractional cointegration analysis.

2. Fractional integration and cointegration

Various definitions of a, fractionally integrated, $I(d)$ process are possible. One asserts that a scalar process a_t, $t \in Z$, is said to be $I(d)$, $d > 0$, if there exists a zero mean scalar $I(0)$ process η_t, $t \in Z$, and a scalar μ, such that

$$a_t = \mu + \Delta^{-d}\eta_t 1(t > 0), \quad t \in Z, \ d > 0, \tag{2}$$

where $1(\cdot)$ is the indicator function $\Delta = 1 - L$, L is the lag operator, and formally

$$(1-z)^d = \frac{1}{\Gamma(-d)} \sum_{j=0}^{\infty} \frac{\Gamma(j-d) z^j}{\Gamma(j+1)}, \qquad \Gamma(\alpha) = \int_0^{\infty} \mathrm{e}^{-x} x^{\alpha-1}\, \mathrm{d}x,$$

results for integer d following from taking $\Gamma(0)/\Gamma(0) = 1$ and then $\Gamma(-d) = \infty$, $d = 0, 1, \ldots$.

In (2), a_t can be said to be asymptotically stationary when $d < \frac{1}{2}$; it is nonstationary solely due to the truncation on the right-hand side. The truncation is designed to cater for cases $d \geqslant \frac{1}{2}$, because otherwise the right-hand side of (2) does not converge in mean square and hence a_t is not well defined; we might refer to a_t given by (2) for $d \geqslant \frac{1}{2}$ as purely nonstationary. An alternative $I(d)$ definition for $d \geqslant \frac{1}{2}$ is

$$a_t = \mu + \Delta^{-k}\zeta_t 1(t > 0), \quad t \in Z, \tag{3}$$

where ζ_t is stationary $I(d-k)$, for the integer k such that $d - \frac{1}{2} < k \leqslant d + \frac{1}{2}$. The distinction between (2) and (3) is discussed by Marinucci and Robinson (1999a); in particular, it is shown there that for $\frac{1}{2} < d < \frac{3}{2}$ first differences of (3) and (2) are asymptotically equivalent in the mean square sense, and consequently it seems difficult to statistically distinguish between them. In this paper, we focus on (2), which we consider more natural and direct as a data generating process; adopting (3) would affect the limiting distribution of some statistics which we use, but not their rates of convergence.

Several definitions relevant to fractional cointegration can be found in the literature (as reviewed by Robinson and Yajima, 2000), but for our purposes it is convenient to elaborate on (1) as follows. We partition z_t as $z_t = (x_t', y_t)'$, where y_t is a scalar and $x_t = (x_{1t}, \ldots, x_{p-1,t})'$. We say that z_t is cointegrated (of orders $d_1, \ldots, d_{p-1}, d_y; d(\beta)$) if x_{it} is $I(d_i)$, $i = 1, \ldots, p-1$, and y_t is $I(d_y)$ and if there exists a $(p-1) \times 1$ vector β such that $e_t = y_t - \beta' x_t$ is $I(d(\beta))$, for $d(\beta) < d_y$. Clearly $d(0) = d_y$, and this definition entails $d_i = d_y > d(\beta)$ for at least one i. In (1) we did not normalize α, but of course if $\alpha' z_t$ is in $I(d(\beta))$ then so is $c\alpha' z_t$ for any $c \neq 0$. However, while the choice of nonzero value for the coefficient of y_t is thence arbitrary, the selection of y_t to have, of necessity, a nonzero coefficient influences the investigation. Alternative normalizations to $\alpha = (-\beta', 1)'$ could provide nontrivially different cointegrating relations, for example $e_t = \beta_1 x_{1t} + \beta_2 x_{2t}$, when $d(\beta) < d_1 = d_2 < d_y$, β_i being the ith element of β. If $d_1 = \cdots = d_{p-1} = d_y$, say, this cannot arise, but our definition reflects the fact that a normalization of a unit type, which is natural in the context of the regression procedures we shall use, requires selection of the normalized variate.

Our cointegration definition implies invariance to inclusion of further variates having integration order no greater than $d(\beta)$. However, the coefficients of these would be unidentified, as indeed are β_i for i such that $d_i \leqslant d(\beta)$. On the other hand, if β is already identified, but we then go on to include further variates that satisfy $d_i \leqslant d(\beta)$, then by partitioned regression it may be shown that the large sample properties described for estimates of β in the following section still hold. Note that when $p \geqslant 3$, the existence of cointegration need not identify even β_i for which $d_i > d(\beta)$. If there is more than one cointegrating relation, so that for some $(p-1)\times 1$ vector $\gamma \neq \beta$, $y_t - \gamma' x_t$ is $I(d(\gamma))$ for $d(\gamma) < d_y$, then it follows that $(\beta - \gamma)' x_t$ is $I(d(\beta,\gamma))$, where $d(\beta,\gamma) \leqslant \max(d(\beta), d(\gamma)) < d_y$, and it may then be shown that there exists no $(p-1)\times(p-1)$ diagonal matrix Λ_n such that $\Lambda_n \sum_{t=1}^n x_t x_t' \Lambda_n$ converges weakly to a matrix that is both finite and nonsingular. However, with $k > 1$ nontrivial, different, cointegrating relations we can redefine x_t as a $(p-k)$ vector and y_t as a $k \times 1$ vector, whence the regression theory referred to in the following section applies to each of the k regressions. Robinson and Yajima (2000) have presented methods for determining fractional cointegrating rank in this situation. Here, we shall proceed in the context of at most one cointegrating relation, and while our stress on the dependence of the integration order of this on β was important to the above discussion, we shall henceforth abbreviate $d(\beta)$ to d_e.

3. Estimation of cointegrating vectors

We discuss estimation of β in the representation

$$y_t = \beta' x_t + e_t \tag{4}$$

for the observable vector $z_t = (x_t', y_t)'$ introduced in the previous section. The setting is that of our definition of cointegration, so that the unobservable process e_t is $I(d_e)$, $d_e < d_y$, while we assume that β is identified. Further to the discussion concluding the previous section, we assume z_t is observed for $t = 1,\ldots,n$.

For a generic column vector or scalar sequence a_t, $t = 1,\ldots,n$, define the discrete Fourier transform

$$w_a(\lambda) = \frac{1}{(2\pi n)^{1/2}} \sum_{t=1}^{n} a_t \mathrm{e}^{\mathrm{i}t\lambda}. \tag{5}$$

With also a column vector or scalar sequence b_t, $t=1,\ldots,n$, possibly identical to a_t, define the (cross-) periodogram

$$I_{ab}(\lambda)=w_a(\lambda)w_b'(-\lambda). \tag{6}$$

Now denote by $\lambda_j=2\pi j/n$, for integer j, the Fourier frequencies, and define the averaged (cross-) periodogram

$$\hat{F}_{ab}(m)=2\,\mathrm{Re}\left\{\frac{2\pi}{n}\sum_{j=1}^{m}I_{ab}(\lambda_j)\right\}-\frac{2\pi}{n}I_{ab}(\pi)1\left(m=\frac{n}{2}\right), \tag{7}$$

where $1(\cdot)$ is the indicator function and the integer m satisfies $1\leqslant m\leqslant n/2$. The last term in (7) only contributes when n is even and m achieves its maximum value, $n/2$. The case $m=[n/2]$, where $[\cdot]$ denotes integer part, is of particular interest, as we deduce that

$$\hat{F}_{ab}\left(\left[\frac{n}{2}\right]\right)=\frac{1}{n}\sum_{t=1}^{n}(a_t-\bar{a})(b_t-\bar{b})', \tag{8}$$

the mean-corrected sample covariance, with $\bar{a}=n^{-1}\sum_{t=1}^{n}a_t$. We observe that $\hat{F}_{ab}(m)$ represent the contributions from frequencies $[1,\lambda_m]$ to the sample covariances in (8); the case where mean-correction is absent from (7)/(8) could also be considered but we omit it due to pressure of space and because it seems less relevant for empirical applications.

We estimate β by the frequency domain least squares (FDLS) statistic

$$\hat{\beta}_m=\hat{F}_{xx}(m)^{-1}\hat{F}_{xy}(m), \tag{9}$$

assuming the inverse exists. Robinson (1994a) proposed (9) when $p=2$, with stationary x_t, y_t in mind. In view of (8), a special case is the ordinary least squares (OLS) estimate with intercept,

$$\hat{\beta}_{[n/2]}=\left(\sum_{t=1}^{n}(x_t-\bar{x})(x_t-\bar{x})'\right)^{-1}\sum_{t=1}^{n}(x_t-\bar{x})(y_t-\bar{y})'. \tag{10}$$

For $m<[n/2]$, there are broadly two cases of interest in the asymptotic context where $n\to\infty$, namely

$$m\sim Cn,\qquad 0<C<\tfrac{1}{2}, \tag{11}$$

and

$$\frac{1}{m}+\frac{m}{n}\to 0. \tag{12}$$

In case (11), a nondegenerate subset of frequencies is used; for this case, $\hat{\beta}_m$ was introduced by Hannan (1963), and subsequently considered by Robinson (1973) and Engle (1974), who named the approach ‘band-spectrum regression’. In case (12), an increasing number of Fourier frequencies is again used, but estimation is carried out over only a degenerating band of frequencies, around the origin.

The OLS estimates (10) have been widely used in AR-based analysis of cointegration for $I(1)$ and $I(2)$ series z_t, initially as estimates of interest in themselves (e.g. Engle and Granger, 1987; Stock, 1987) and latterly as initial estimates used to compute residuals which are then employed in producing estimates of β with superior properties (e.g. Phillips, 1991a, b; Phillips and Hansen, 1990). The classical regression assumption of orthogonality between e_t and x_t in (3) is not imposed, but nevertheless $\hat{\beta}_m$ is still consistent because of the asymptotic dominance of e_t by x_t.

This lack of orthogonality results, however, in loss of consistency of least squares when x_t is stationary (so $d_i < \frac{1}{2}$, all i), the usual simultaneous equations bias resulting. This motivated the first consideration of $\hat{\beta}_m$ under (12) in the cointegration setting, by Robinson (1994a), who showed that despite correlation between the stationary x_t and e_t, $\hat{\beta}_m$ is consistent for β due to the dominance of the spectrum of x_t over that of e_t near zero frequency (this is not the case under (11)). Thus $\hat{\beta}_m$ under (12) is superior to OLS (10) for stationary observables. In practice, one may not be sure whether or not observables are stationary, especially in a fractional context where the transition between stationarity and nonstationarity is smooth. This partly motivated Robinson and Marinucci's (1997, 1999, 2000) theoretical study of $\hat{\beta}_m$ for purely nonstationary $I(d)$ z_t, another motivation being that, since cointegration as discussed above is essentially a low frequency phenomenon, inclusion of high frequency contributions might seem unwarranted. Indeed, a major technical focus of this work was avoidance of assumptions on behaviour at frequencies away from zero, so that the input process (see η_t in (2)) is $I(0)$ in the very general sense described at the start of Section 1.

The properties of $\hat{\beta}_m$ depend on the integration orders $d_1, \ldots, d_{p-1}$ and d_e, and they also vary in a more qualitative fashion between several regions of the $d_1, \ldots, d_{p-1}, d_e$-space. To keep the description simple we omit detailed regularity conditions and suppose that each of the $d_1, \ldots, d_{p-1}$ satisfies the same restrictions; in fact the discussion of Robinson and Marinucci (1999a), where the main focus was the properties of the averaged periodogram itself, was confined to $p=2$. We denote by $\hat{\beta}_{im}$ the ith element of $\hat{\beta}_m$, and assume (12) holds throughout. We report only rates of convergence, taking $X_n \backsim_d h(n)$ to mean that $h(n)^{-1}X_n$ converges in distribution to a well-defined, nondegenerate random variable, so that this implies $X_n = O_p(h(n))$. For $i = 1, \ldots, p-1$, we have the following cases:

(i) ('Less than unit root nonstationarity'): $d_i > \frac{1}{2}$, $d_e \geqslant 0$, $d_i + d_e < 1$.

$$\hat{\beta}_{i[n/2]} - \beta_i \backsim_d n^{1-d_i-d_{\min}}, \qquad \hat{\beta}_{im} - \beta_i \backsim_d n^{d_e-d_i} m^{1-d_{\min}-d_e}. \tag{13}$$

(ii) ('Boundary case'): $\frac{1}{2} \leqslant d_i = 1 - d_e < 1$.

$$\hat{\beta}_{i[n/2]} - \beta_i \backsim_d n^{2d_e-1}\log n, \qquad \hat{\beta}_{im} - \beta_i \backsim_d n^{2d_e-1}\log m. \tag{14}$$

(iii) ('$I(1)/I(0)$ case'): $d_1 = \cdots = d_{p-1} = 1,\ d_e = 0$.

$$\hat{\beta}_{i[n/2]} - \beta_i \backsim_d n^{-1}, \qquad \hat{\beta}_{im} - \beta_i \backsim_d n^{-1}. \tag{15}$$

(iv) ('Greater than unit root nonstationarity'): $d_1 = \cdots = d_{p-1} > \frac{1}{2},\ d_e > 0$, $d_i + d_e > 1$.

$$\hat{\beta}_{i[n/2]} - \beta_i \backsim_d n^{d_e-d_i}, \qquad \hat{\beta}_{im} - \beta_i \backsim_d n^{d_e-d_i}. \tag{16}$$

The limit random variables here all have nonstandard distributions, further details of which can be found in Robinson and Marinucci (1999), who employ functional limit theory of Marinucci and Robinson (2000). Case (iii) is the usual $I(1)/I(0)$ one from the AR-based cointegration literature, and our presentation obscures the fact that FDLS enjoys some 'second-order bias' superiority over OLS; this case is discussed in greater detail by Marinucci and Robinson (1999b). (14) indicates that the degenerate FDLS can converge slightly faster, or have a more concentrated limit distribution, than OLS, while case (i) demonstrates a clear-cut superiority in the former approach; (16) shows that so long as at least an arbitrarily slowly increasing number m of frequencies is included, omission of higher frequencies makes no difference at all to limit distributional behaviour when the collective memory of the regressor and the cointegrating error e_t exceeds 1. The reason why omission of frequencies causes no damage, and even some improvement, is that on the one hand it eliminates the simultaneous equation bias due to the omitted frequencies, while on the other, variance is dominated by contributions from low frequencies, due to nonstationarity.

We shall not be concerned with the possibility of deterministic components, a situation discussed by Robinson and Marinucci (2000): the results described under cases (i)–(iv) continue to hold when the deterministic trends are effectively dominated by the stochastic ones, whereas if the deterministic trends dominate then $\hat{\beta}_{im}$ is asymptotically normal, to contrast with the nonstandard limit laws in (13)–(16).

4. Statistical inference on integration orders

In view of Definition 2, inference on integration orders $d_1, \ldots, d_p, d_e$ is bound to be a key part of any investigation of fractional cointegration. As with estimation of β, semiparametric methods based on only a degenerating band of low frequencies are stressed. Semiparametric estimates are robust in that they achieve consistency without the need for a parametric model, misspecification of which can cause inconsistency of estimates of integration orders.

Our fractional definitions in Section 2 describe only scalar sequences, whereas we will sometimes be concerned with inference on integration orders for a vector process, and this requires us to think in terms of a model for jointly dependent fractional processes. Estimation of integration orders has principally been developed under covariance stationarity assumptions, which allow a spectral density to exist. We will discuss the topic in this setting, because the same asymptotic statistical properties are likely to hold for asymptotically stationary processes (i.e. $d < \frac{1}{2}$), while purely nonstationary processes (such that $d \geqslant \frac{1}{2}$) might be handled by integer differencing to produce at least asymptotic covariance stationarity, whereupon the methods of integration order estimation which we discuss, justified for stationary series by Robinson (1995a, b), Lobato (1999), can be applied, and then the order of integer differencing added back. Notice that whereas in an AR setting inference based on differenced data can be very inefficient, this is not the case in a fractional setting, where indeed rules that are efficient in the classical sense are based on the null differenced data, see Robinson (1994b). Also, first differencing caters automatically for a linear trend as it reduces it to a constant, to which the discrete Fourier transform (5) evaluated at $\lambda = 2\pi j/n$, $j = 1, \ldots, n/2$ is invariant. To estimate the integration order d_e of e_t we can apply our techniques with the generic process ξ_t discussed below representing $\hat{e}_t = y_t - \hat{\beta}' x_t$, where $\hat{\beta}$ is one of the estimates described in Section 3.

Consider a $q \times 1$ covariance stationary vector process ξ_t, $t = 0, \pm 1, \ldots$, having spectral density matrix $f(\lambda)$, whose (k, ℓ)th element $f_{k\ell}(\lambda)$ satisfies

$$f_{k\ell}(\lambda) \sim g_{k\ell} e^{i(\pi/2)(\delta_k - \delta_\ell)} \lambda^{-\delta_k - \delta_\ell} \quad \text{as } \lambda \to 0^+ \tag{17}$$

for $k, \ell = 1, \ldots, q$, with $\sim$ indicating that the ratio of real parts of left- and right-hand sides, and the ratio of imaginary parts of left- and right-hand sides, both tend to 1, and $-\frac{1}{2} < \delta_k < \frac{1}{2}$, $k = 1, \ldots, q$. The matrix $G = (g_{k\ell})$ is positive definite if ξ_t is not cointegrated, and positive semidefinite otherwise; in any case $g_{kk} > 0$, $k = 1, \ldots, q$.

Two basic approaches to estimation of $\delta = (\delta_1, \ldots, \delta_q)'$ will be employed. The first is the log periodogram estimate of Geweke and Porter-Hudak (1983), justified asymptotically by Robinson (1995a) and Hurvich et al. (1998). Denote by $I_{kk}(\lambda)$ the kth diagonal element of $I_{\xi\xi}(\lambda)$, (see (6)). For integer s, define $Y_{kj} = \log(I_{kk}(\lambda_j))$, $k = 1, \ldots, q$, $j = 1, \ldots, s$, where $s < [n/2]$; s is a bandwidth parameter, somewhat analogous to m introduced in the previous section, but it must tend to ∞ with n no faster than a rate determined by the smoothness of the functions $f_{kk}(\lambda)\lambda^{2\delta_k}$ at $\lambda = 0$. Define

$$\tilde{\delta}_k = -\frac{\sum_{j=1}^{s} v_j Y_{kj}}{2\sum_{j=1}^{s} v_j^2}, \quad v_j = \log j - s^{-1} \sum_{j=1}^{s} \log j, \;\; k = 1, \ldots, q. \tag{18}$$

We have the approximations $\tilde{\delta}_k \sim N(\delta_k, \pi^2/24s)$, $k=1,\ldots,q$ (Robinson, 1995a); a Wald-type test of the hypothesis $H_0: \Pi\delta = \rho$, for a prescribed $u \times q$ matrix Π and $u \times 1$ vector ρ, is given by rejecting the null if

$$\tilde{W} = 4s(\Pi\tilde{\delta} - \rho)'(\Pi\tilde{\Omega}\Pi')^{-1}(\Pi\tilde{\delta} - \rho) \tag{19}$$

is significantly large relative to the χ^2_u distribution, where $\tilde{\delta} = (\tilde{\delta}_1, \ldots, \tilde{\delta}_q)'$ and $\tilde{\Omega}$, a consistent estimate of the limiting variance matrix of $2s^{1/2}(\tilde{\delta} - \delta)$, is defined by Robinson (1995a). Note that these and other semiparametric estimates of integration orders are less-than-$n^{1/2}$-consistent (by contrast with estimates of β from nonstationary data), so that reliable inference on the δ_i requires a large enough sample.

In case restrictions in the δ_k are detected, more efficient estimates are available, again on the lines of Robinson (1995a). Assume it has been established that $\delta_1 = \cdots = \delta_q$, and we wish to estimate the common value δ_*. We consider the GLS-type estimate

$$\tilde{\delta}_* = -\frac{\sum_{j=1}^{s} 1_q'\tilde{\Omega}^{-1} Y_j v_j}{2 1_q'\tilde{\Omega}^{-1} 1_q \sum_{j=1}^{s} v_j^2}, \tag{20}$$

where $Y_j = (Y_{1j}, \ldots, Y_{qj})'$. We can use the approximation $\tilde{\delta}_* \sim N(\delta_*, (1_q'\tilde{\Omega}^{-1}1_q)^{-1}/4s)$, so a Wald test that δ_* takes on a particular value can readily be conducted.

The efficiency of $\tilde{\delta}, \tilde{\delta}_*$ is inferior to another class of semiparametric estimates, the narrow-band Gaussian or Whittle estimate, introduced by Künsch (1987) and developed by Robinson (1995b) and Lobato (1999). This essentially optimizes an approximate form of Gaussian likelihood, but extending only over the s smallest Fourier frequencies λ_j. Consider first the q individual univariate objective functions

$$R_k(\delta_k) = \log\left(\frac{1}{s}\sum_{j=1}^{s} I_{kk}(\lambda_j) j^{2\delta_k}\right) - \frac{2\delta_k}{s}\sum_{j=1}^{s}\log j, \tag{21}$$

and estimates $\bar{\delta}_k = \arg\min(R_k(\delta_k))$, for $k=1,\ldots,q$, minimizing over a suitable compact subset of $(-\frac{1}{2}, \frac{1}{2})$, to impose stationarity and invertibility. Then, individually, we have the approximation $\bar{\delta}_k \sim N(\delta_k, (4s)^{-1})$, $k=1,\ldots,q$, so the $\bar{\delta}_k$ are more efficient than the $\tilde{\delta}_k$ (Robinson, 1995b). A further efficiency improvement, when $q>1$, follows from the multivariate objective function (Lobato, 1996)

$$R(\delta) = \log\left|\frac{1}{s}\sum_{j=1}^{s} A_j(\delta)\right| - \frac{2}{s}\sum_{j=1}^{q}\delta_k\sum_{j=1}^{s}\log j, \tag{22}$$

where $A_j(\delta) = \mathrm{Re}\{\Lambda_j(\delta) I_{\xi\xi}(\lambda_j) \bar{\Lambda}_j(\delta)\}$, $\Lambda_j(\delta) = diag\{e^{i\pi\delta_1/2} j^{\delta_1}, \ldots, e^{i\pi\delta_q/2} j^{\delta_q}\}$. Then define $\hat{\delta} = \arg\min R(\delta)$, minimizing over a compact subset of $(-\frac{1}{2}, \frac{1}{2})^q$. Unlike $\tilde{\delta}$ and $\tilde{\delta}_*$, the $\bar{\delta}_k$ and $\hat{\delta}$ are not defined in closed form. However, commencing from an initial $s^{1/2}$-consistent estimate, an estimate with identical asymptotic efficiency is achieved by a single approximate Newton step; further such steps offer no further first order efficiency improvement, though they may improve higher order efficiency (see Robinson, 1988). Considering only $\hat{\delta}$, to describe the $(v+1)$th step in such a procedure, $v \geqslant 0$, denote by $\hat{\delta}^{[v]}$ the current estimate and

$$\hat{\delta}^{[v+1]} = \hat{\delta}^{[v]} - \{2(I_q + \hat{G}^{[v]} \circ \hat{G}^{[v]^{-1}})\}^{-1} \frac{\partial R(\hat{\delta}^{[v]})}{\partial \delta}, \tag{23}$$

where '$\circ$' denotes Hadamard product and $\hat{G}^{[v]} = \hat{G}(\hat{\delta}^{[v]})$, where $\hat{G}(\delta) = s^{-1} \sum_{j=1}^{s} A_j(\delta)$. The matrix in braces in (23) is a simple consistent estimate of the probability limit of $\partial^2 R(\hat{\delta}^{[v]})/\partial\delta\partial\delta'$; note that $2(I_q + \hat{G}^{[v]} \circ \hat{G}^{[v]^{-1}})$ reduces to 4 when $q = 1$. A possible choice for $\hat{\delta}^{[0]}$ is $\tilde{\delta}$, since the latter estimate is like $\hat{\delta}$, $s^{1/2}$-consistent, though it is less efficient. We stress that the desirable properties described above assume no cointegration when ξ_t is a vector, but as before apply to individual elements of ξ_t when it is cointegrated.

A final efficiency improvement is available if any a priori restrictions are incorporated. Considering again the case $\delta_1 = \cdots = \delta_q = \delta_*$, with δ_* unknown, define

$$R_*(\delta) = \log\left|\frac{1}{s}\sum_{j=1}^{s} A_j(\delta_* 1_q)\right| - \frac{2q}{s}\sum_{j=1}^{s}\log j, \tag{24}$$

and $\hat{\delta}_* = \arg\min R_*(\delta)$. Equal efficiency is achieved by the estimates

$$\hat{\delta}_*^{[v+1]} = \hat{\delta}_*^{[v]} - \frac{1}{4q}\frac{\partial R(\hat{\delta}^{[v]} 1_q)}{\partial \delta}, \quad v \geqslant 0, \tag{25}$$

given an initial $s^{1/2}$-consistent $\hat{\delta}_*^{[0]}$ (such as $\hat{\delta}_*$).

Wald tests based on $\bar{\delta}, \hat{\delta}, \hat{\delta}_*$ and the corresponding Newton steps are available; for example, under the null hypothesis $\Pi\delta = \rho$, the statistic

$$\hat{W}_{\Pi\delta=\rho} = (\Pi\hat{\delta} - \rho)'\{2\Pi(I_q + \hat{G}(\hat{\delta}) \circ \hat{G}(\hat{\delta})^{-1})\Pi'\}(\Pi\tilde{\delta} - \rho), \tag{26}$$

has a limiting χ^2_u distribution. Likewise, a test on δ_* can be based on the approximation $\hat{\delta}_* \sim N(\delta_*, (4qs)^{-1})$. The objective functions (21), (22) and (24) suggest also the use of tests based on the Lagrange Multiplier (LM) and Likelihood Ratio (LR) principles. In fact, an LM-type test of $\delta_1 = \cdots = \delta_q = 0$, against a somewhat different alternative than (17), was proposed by Lobato and Robinson (1998), while an LR type test in case $q=1$ was proposed by Robinson (1998). In our context, an LM statistic for testing $\Pi\delta=\rho$ is

$$LM_{\Pi\delta=\rho}=s\frac{\partial R(\hat{\delta}_0)}{\partial\delta'}\{2\Pi(I_q+\hat{G}(\hat{\delta}_0)\circ\hat{G}(\hat{\delta}_0)^{-1})\Pi'\}\frac{\partial R(\hat{\delta}_0)}{\partial\delta}, \tag{27}$$

where $\hat{\delta}_0$ minimizes $R(\delta)$ subject to $\Pi\delta=\rho$, or else is a Newton approximation to this, computed along the lines described above. Then, from asymptotic theory of Robinson (1995b), Lobato (1996), (27) has a limiting null χ^2_u distribution. To test $H_0:\delta_*=\delta_{*0}$ an LM statistic is

$$LM_{\delta_*=\delta_{*0}}=s\left\{\frac{\partial R_*(\delta_0)}{\partial\delta}\right\}^2\Bigg/4q, \tag{28}$$

and has a limiting null χ^2_1 distribution. LR-type statistics for testing $H_0:\Pi\delta=\rho$ and $H_0:\delta_*=\delta_{*0}$ are, respectively,

$$LR_{\Pi\delta=\rho}=2s\{R(\hat{\delta}_0)-R(\hat{\delta})\},\quad LR_{\delta_*=\delta_{*0}}=2s\{R_*(\hat{\delta}_{*0})-R(\hat{\delta}_*)\}, \tag{29}$$

and have asymptotic null χ^2_u and χ^2_1 distributions, respectively.

We now consider the problem of testing for the presence, or absence, of cointegration, given that we have established from the procedures described above that at least two observables have the same integration order. Robinson and Yajima (2000) consider an approach for determining fractional cointegrating rank in a general vector context that requires introduction of an additional user-chosen tuning number. Our present approach does not but, for simplicity of exposition, and because it suffices for the empirical examples of the following section, we focus on a bivariate observable, so $p=2$, and at most one cointegrating relation can exist. We present a test for the null hypothesis of no cointegration based on the same principle as that employed by Hausman (1978) in other settings. We do not provide rigorous theoretical support, but in addition to using it in our empirical examples we will report Monte Carlo experiments that investigate its validity and power. Consider again the set-up of Section 4, with $p=2$. Focusing on the Gaussian approach, under the necessary condition $\delta_1=\delta_2=\delta_*$ for cointegration, with δ_* unspecified, recall that the univariate estimates $(\bar{\delta}_1,\bar{\delta}_2)$ consistently estimate $(\delta_1,\delta_2)=(\delta_*,\delta_*)$. However, both $\bar{\delta}_1$ and $\bar{\delta}_2$ are less efficient asymptotically than the restricted estimate $\hat{\delta}_*$ when $\delta_1=\delta_2$ and G is positive definite, so there is no cointegration. On the other hand, if ξ_t is cointegrated, it appears that $\hat{\delta}_*$ is inconsistent for δ_*; the original Gaussian objective function is not well defined when G is singular, and so there is no basis for considering the concentrated form (24) as an objective function in the first place. We can thus test for no cointegration indirectly, comparing $\hat{\delta}_*$ with, say, $\bar{\delta}_1$. Because $\bar{\delta}_1$ has asymptotic variance $1/4s$, while $\hat{\delta}_*$ has asymptotic variance $1/8s$ under $d=d_e$, it follows by an argument along the lines of Hausman (1978) that

$\hat{\delta}_* - \bar{\delta}_1$ has asymptotic variance $1/4s - 1/8s = 1/8s$, and then proceeding as in Robinson (1995b),

$$H_{1s} = 8s(\hat{\delta}_* - \bar{\delta}_1)^2 \to_d \chi_1^2 \quad \text{as } \frac{1}{s} + \frac{s}{n} \to 0. \tag{30}$$

This argument is heuristic but seems sufficiently convincing for the test to warrant serious consideration.

5. Empirical examples

Our empirical work employs macroeconomic data of Engle and Granger (1987) (consumption and income, quarterly data, 1947Q1-1981Q2) and Campbell and Shiller (1987) (stock prices and dividends, annual data, 1871–1986). For each bivariate series, denote by y the variable chosen to be ‘dependent’ and by x the ‘independent’ one in our definition of cointegration, and by d_y, d_x the respective integration orders. Our results, based on the methodology described in Sections 3 and 4, are presented in three steps.

5.1. Memory of raw data

We estimated d_x, d_y by supposing that both lie between $\frac{1}{2}$ and $\frac{3}{2}$, first-differencing the x and y series, applying procedures of Section 4 to estimate $\delta_x = d_x - 1$ and $\delta_y = d_y - 1$, and then adding 1. As motivated in Section 4, we computed univariate log-periodogram $(\tilde{\delta}_x, \tilde{\delta}_y)$ and Gaussian $(\bar{d}_x, \bar{d}_y)$ estimates; bivariate, unrestricted Gaussian estimates $\hat{d}_x, \hat{d}_y$; bivariate, log-periodogram $(\tilde{d}_*)$ and Gaussian $(\hat{d}_*)$ estimates of a common $d_x = d_y$.

We report the estimates in Tables 1 and 2. For all but the univariate log-periodogram estimates we report also approximate 95% confidence intervals (denoted *CI* in the tables) based on the (normal) asymptotic distribution theory developed by Robinson (1995a, b), Lobato (1996, 1999). In order to judge sensitivity to choice of bandwidth s, we chose a grid of three values for each data set analysed.

Gaussian estimates were approximated by the Newton steps described in Section 4, iterating until convergence to 5 decimal places. In the univariate and constrained cases, the objective function is globally concave, as shown for the univariate case by application of the Cauchy inequality to (4.3) of Robinson (1995b), and the version of Newton iteration employed guarantees eventual convergence in such circumstances. On the other hand, for one data set (consumption and income) we found lack of convergence of the unconstrained bivariate Newton step procedures for the smallest s, $s = 22$; this may be due to the presence of cointegration, as argued in Section 4, or could instead be due to poor model fit. We omit the corresponding results from

Table 1
Consumption (C) and income (I) (Eqs. (18, 20, 23, 25))

	s	$\tilde{d}$	$\bar{d}$	CI	$\hat{d}$	CI	$\tilde{d}_*$	CI	$\hat{d}_*$	CI
C	22	1.13	1.13	0.92, 1.34	—	0.67, 1.04	0.95	0.94, 0.97	0.89	0.74, 1.03
	30	1.04	1.11	0.93, 1.29	0.97	0.81, 1.13	0.98	0.97, 0.99	0.96	0.83, 1.08
	40	1.04	1.12	0.97, 1.28	1.00	0.86, 1.14	1.02	1.02, 1.03	1.02	0.91, 1.13
I	22	0.89	0.99	0.78, 1.20	—	0.74, 1.11	0.95	0.94, 0.97	0.89	0.74, 1.03
	30	0.95	1.03	0.85, 1.21	0.96	0.80, 1.12	0.98	0.97, 0.99	0.96	0.83, 1.08
	40	1.02	1.08	0.92, 1.23	1.03	0.89, 1.17	1.02	1.02, 1.03	1.02	0.91, 1.13

Table 2
Stock prices (S) and dividends (D) (Eqs. (18, 20, 23, 25))

	s	$\tilde{d}$	$\bar{d}$	CI	$\hat{d}$	CI	$\tilde{d}_*$	CI	$\hat{d}_*$	CI
S	22	0.96	1.04	0.83, 1.25	0.89	0.71, 1.06	0.94	0.92, 0.96	0.88	0.73, 1.03
	30	0.83	0.91	0.73, 1.09	0.78	0.62, 0.94	0.84	0.83, 0.85	0.77	0.65, 0.90
	40	0.84	0.90	0.75, 1.06	0.82	0.68, 0.96	0.87	0.86, 0.87	0.82	0.71, 0.93
D	22	0.91	0.88	0.67, 1.09	0.84	0.67, 1.02	0.94	0.92, 0.96	0.88	0.73, 1.03
	30	0.86	0.86	0.68, 1.03	0.75	0.59, 0.91	0.84	0.83, 0.85	0.77	0.65, 0.90
	40	0.91	0.95	0.80, 1.11	0.83	0.70, 0.97	0.87	0.86, 0.87	0.82	0.71, 0.93

Table 1, and ones dependent on those cases in subsequent tables. The univariate estimates are consistent both under cointegration and under no-cointegration, and they tend to produce very close estimates for a given s; confidence intervals always refer to estimates to their immediate left.

Although only $d_x = d_y$ is necessary for cointegration in our fractional approach, to comply with tradition we first test $d_x = d_y = 1$, that is $\delta_x = \delta_y = 0$; note that this case occupies a set of measure zero in the fractional domain. We employ (see Table 3) $\hat{W}_{\Pi\delta=\rho}$ and $LM_{\Pi\delta=\rho}$, with $\Pi = I_2$ and $\rho = (0,0)'$, so that $\hat{\delta} = (0,0)'$ in (27), denoting this version of (27) by $LM^{(2)}$ in the tables, as well as $LR_{\delta_* = \delta_{*0}}$ with $\delta_{*0} = 0$. Because these procedures, based on the bivariate series, are liable to be invalid under cointegration, we also used univariate versions of $LM_{\Pi\delta=\rho}$, with $\Pi = 1$, $\rho = 0$, on each individual series y_t, x_t, denoting the statistics $LM_y^{(1)}$, $LM_x^{(1)}$, respectively. The 5% and 1% χ_2^2 critical values for $W_{\Pi\delta=\rho}$ and $LM^{(2)}$ are 5.99 and 9.21, and the 5% and 1% critical values for $LR_{\delta_* = \delta_{*0}}$, $LM_y^{(1)}$ and $LM_x^{(1)}$ are 3.84 and 6.63. We used the same values of s as in Tables 1–2.

Globally, we fail to reject the $I(1)$ null in all 15 cases for consumption and income; the evidence provided by stock prices and dividends is mixed, with four rejections out of 15 at 5%.

We now abandon the unit root null to focus on the restriction $d_x = d_y$, whose value is unknown under the null. We report in Table 4 the statistics $\tilde{W}$, $\hat{W}_{\Pi\delta=\rho}$, $LR_{\Pi\delta=\rho}$ and $LM_{\Pi\delta=\rho}$, where $\Pi = (1, -1)'$, $\rho = 0$, comparing with χ_1^2 critical values.

Since the null $d_x = d_y$ nests the joint $I(1)$ assumption, it is not surprising that we fail to reject in all 12 cases for consumption and income; there is also evidence that stock prices and dividends share the same, possibly nonunit, integration order.

5.2. Cointegrating regression estimates and diagnostics

Table 5 contains OLS as well as, for three other values of m, FDLS; the values of m used (3, 4 and 6) are much smaller than the bandwidths s used in inference on d_x and d_y due to the anticipation of nonstationarity in the raw data; for stationary x_t, y_t optimal rules of bandwidth choice would lead to m that are more comparable with the s we have used. For each m, we report also the fractions

$$r_{xx,m} = \frac{\hat{F}_{xx}(1,m)}{\hat{F}_{xx}(1,[(n-2)/2])}, \qquad r_{xy,m} = \frac{\hat{F}_{xy}(1,m)}{\hat{F}_{xy}(1,[(n-2)/2])}, \tag{31}$$

their closeness to unity being an indicator of support for the basic rationale behind our approach, that sample variability eventually concentrates at zero frequency for nonstationary processes (though note that $r_{xy,m}$ need not lie in $[0,1]$.)

5.3. Memory of cointegrating error

We computed several estimates of d_e. In Table 6, we report $\tilde{d}_e$ and $\hat{d}_e$, the log-periodogram and Gaussian estimates of (i) and (iii) above, using first differences of the residuals and then adding unity. However, in case there is cointegration, nonstationary e_t seem rather unlikely a priori in our series, and so we also report, in Table 7, corresponding estimates of d_e based simply on raw data and without addition. Also, we report 95% confidence intervals based on the asymptotic theory of Robinson (1995a, b), though strictly this has not been justified in case of the residuals $\hat{e}_t$; we conjecture that asymptotic distributions are unaffected by the presence of estimated parameters, at least when FDLS are more-than-$n^{1/2}$-consistent, as is often the case under nonstationarity, see (13)/(16).

Tests were also conducted in order to more directly investigate the possibility of cointegration. We begin by again catering to the reader schooled in traditional cointegration analysis by testing $I(1)$ and $I(0)$ hypotheses, albeit against fractional alternatives. Table 8 reports univariate LM statistics (27) based on residuals for testing $d_e = 1$ and $d_e = 0$, these being equivalent

Table 3
Tests for unit roots (Eqs. (26, 29, 27))

	$W_{\Pi\delta=\rho}$		$LR_{\delta_*=\delta_{*0}}$				$LM^{(2)}$			$LM_y^{(1)}$			$LM_x^{(1)}$		
	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$
C/I	—	0.37	0.23	1.4	0.26	0.05	1.1	0.37	0.10	0.39	0.56	1.2	0.00	0.05	0.61
S/D	3.4	13.3	9.6	0.98	6.5	5.4	0.78	3.2	2.9	0.05	0.49	0.77	0.14	0.50	0.09

Table 4
Tests for $d_y = d_x$ ($\delta_y = \delta_x$) (Eqs. (19, 26, 27))

	$\tilde{W}$			$\hat{W}_{\Pi\delta=\rho}$			$LR_{\Pi\delta=\rho}$			$LM_{\Pi\delta=\rho}$		
	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$
C/I	1.06	0.02	0.02	—	0.07	0.66	—	0.04	0.28	0.10	0.01	0.11
S/D	0.07	0.04	0.36	0.75	0.43	0.12	0.33	0.19	0.05	0.03	0.04	0.01

Table 5
Cointegration analysis (Eqs. (9, 13))

	$\hat{\beta}_{m_1}$	$\hat{\beta}_{m_2}$	$\hat{\beta}_{m_3}$	$\hat{\beta}_{[n/2]}$	r_{xx,m_1}	r_{xx,m_2}	r_{xx,m_3}	r_{xy,m_1}	r_{xy,m_2}	r_{xy,m_3}
C/I	0.231	0.232	0.232	0.229	0.85	0.88	0.93	0.86	0.88	0.93
S/D	33.2	33.6	32.5	31.0	0.78	0.79	0.85	0.84	0.85	0.89

Table 6
Log-periodogram estimates of d_e (Eq. (18))

	Differences of residuals						Raw residuals					
	$s=22$		$s=30$		$s=40$		$s=22$		$s=30$		$s=40$	
	$\tilde{d}_e$	CI	$\tilde{d}_e$	CI	$\tilde{d}_e$	CI	$\tilde{d}_e$	CI	$\tilde{d}_e$	CI	$\tilde{d}_e$	CI
C/I	0.56	0.39, 0.74	0.84	0.58, 1.1	0.86	0.66, 1.06	0.19	−0.07, 0.46	0.57	0.26, 0.87	0.61	0.38, 0.84
S/D	0.72	0.35, 1.1	0.57	0.28, 0.87	0.64	0.40, 0.88	0.74	0.47, 1.01	0.60	0.38, 0.82	0.64	0.45, 0.83

Table 7
Gaussian estimates of d_e (Eq. (23))

	Differences of residuals						Raw residuals					
	$s=22$		$s=30$		$s=40$		$s=22$		$s=30$		$s=40$	
	$\bar{d}_e$	CI	$\bar{d}_e$	CI	$\bar{d}_e$	CI	$\bar{d}_e$	CI	$\bar{d}_e$	CI	$\bar{d}_e$	CI
C/I	0.62	0.41, 0.83	0.78	0.60, 0.96	0.87	0.71, 1.02	0.44	0.23, 0.65	0.68	0.50, 0.86	0.76	0.61, 0.91
S/D	0.77	0.56, 0.98	0.62	0.44, 0.80	0.65	0.50, 0.81	0.77	0.56, 0.98	0.62	0.44, 0.80	0.62	0.46, 0.77

Table 8
Testing for (no) cointegration (Eqs. (27, 30))

	LM on $d_e=1$			LM on $d_e=0$			H_{ys}			H_{xs}		
	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$	$s=22$	$s=30$	$s=40$
C/I	2.23	0.96	0.61	4.54	14.90	33.84	10.79	5.42	3.55	2.00	1.12	1.30
S/D	0.75	4.57	4.76	16.51	30.68	55.10	4.61	4.37	1.96	0.00	1.56	5.21

respectively to no-cointegration and cointegration in an AR set-up. It is notable, then, that the $I(0)$ null for e_t is rejected on all occasions. Finally, the Hausman test of Section 4 was also employed. Because our stress on testing estimates of δ_1 in Section 4 was arbitrary, we report in Table 8 not only H_{xs} but also H_{ys}, see (30).

We now discuss the implications of the tables for our two pairs of empirical series.

(a) For consumption (y) and income (x), Engle and Granger (1987) found evidence of $CI(1)$ cointegration. Table 1 suggests an integration order very close to one for both variables, the estimates ranging from 0.89 to 1.08 for income and from 0.89 to 1.13 for consumption. The hypothesis $d_x = d_y$ can safely not be rejected as the test statistics are at most 1.06. The $\hat{\beta}_m$ are about 0.232, which is close to OLS (0.229). Variability concentrates rapidly around frequency zero, 85.1% of the variance of income being accounted for by the three smallest periodogram ordinates, less than 5% of the total. This proportion rises to 92.6% for $m=6$ frequencies, and is even greater for the cross-periodogram, confirming the high coherency of the two series at low frequencies. The residual diagnostics are less clear-cut, but in only one case out of 12 does the confidence interval for d_e include zero, providing strong evidence against weak dependence; likewise, the LM test for $I(0)$ is always significant at 5%. The estimates of d_e vary quite noticeably with s and the procedure adopted, ranging from 0.19 to 0.87. The Hausman test for no cointegration rejects in two out of six cases. Overall, it seems the $I(1)/I(0)$ framework can produce a satisfactory approximation for the behaviour of the raw series, but not so for cointegrating residuals. Note that these data are seasonally adjusted, but because in this paper we use only local-to-zero frequency assumptions on the behaviour of the spectral density, seasonal adjustment procedures have no effect asymptotically.

(b) The idea that stock prices (y) and dividends (x) might be cointegrated follows mainly from a present value model, which asserts that an asset price is linear in the present discounted value of future dividends,

$y_t = \theta(1-\delta)\sum_{i=0}^{\infty}\delta^i E_t(x_{t+i}) + c$, where δ is the discount factor; see Campbell and Shiller (1987). In Table 2, the estimates of d_x, d_y appear close to unity, although now the hypothesis that dividends are mean-reverting ($d_x < 1$) appears to be supported. The statistics for testing $d_x = d_y$ are always manifestly insignificant; the evidence on the unit root assumption is more ambiguous, with 4 rejections out of 12 cases. Empirical evidence of cointegration is weak; the estimates of d_e range from 0.57 to 0.77; the Hausman test of no cointegration rejects in three out of six cases. The results of Campbell and Shiller on this data set were, in their own words, inconclusive; our findings are possibly closer to those of Phillips and Ouliaris (1988), who were unable to reject the null of no cointegration at the 10% level.

6. Monte Carlo evidence

To compare the performance of versions of FDLS with OLS in moderate sample sizes a small Monte Carlo study was conducted. Let $u_t = (u_{1t}, u_{2t})'$ be a sequence of independent bivariate normal variates such that u_{1t} and u_{2t} have zero mean, unit variance, and correlation 0.5. We consider the model

$$y_t = \beta x_t + e_t, \quad \beta = 2, \tag{32}$$

$$x_t = \Delta^{-d_x}\{u_{1t}1(t>0)\}, \qquad e_t = \Delta^{-d_e}\{u_{2t}1(t>0)\}. \tag{33}$$

In terms of cases (i)–(iv) of Section 3, we include three versions of case (i) $((d_e, d_x) = (0, 0.8), (0.2, 0.5), (0.4, 0.5))$, one of case (ii) $((d_e, d_x) = (0.2, 0.8))$ and four of case (iv) $((d_e, d_x) = (0, 1.2), (0.2, 1.2), (0.4, 0.8), (0.4, 1.2))$, along with $(d_e, d_x) = (0, 0.5)$ which is excluded from the discussion of Section 3 but is of interest as x_t is on the stationary/nonstationary boundary. We excluded case (iii) as this was examined in detail by Marinucci and Robinson (1999b).

We generated series of lengths $n = 64, 128$, for each of which we computed $\hat{\beta}_m$ for three values of m: $(m_1, m_2, m_3) = (3, 4, 5)$, $(3, 4, 6)$ and $(6, 8, 10)$, respectively, as well as $\hat{\beta}_{[n/2]}$. Tables 9 and 10 report Monte Carlo bias and mean squared error (MSE) across 5000 replications. OLS is always beaten by FDLS; to bear out remarks in Section 3, the relative improvement over OLS seems greatest overall in cases (i) and (ii), while mostly the choice of m in FDLS seems to make little difference, though it is striking that $\hat{\beta}_m$, using the smallest m, is often best even in MSE. It is important to stress that FDLS is unlikely to be the best approach for estimating β. In $I(1)/I(0)$ cointegration analysis, fully modified and system estimates can lead to greater efficiency and standard asymptotics (Phillips, 1991a, b; Phillips and Hansen, 1990), but while work is underway to extend such ideas to our setting this is

Table 9
Bias for FDLS and OLS (Eqs. (9, 10))

n	$d_x = 0.5$					$d_x = 0.8$				$d_x = 1.2$			
	d_e	$\hat{\beta}_{m_1}$	$\hat{\beta}_{m_2}$	$\hat{\beta}_{m_3}$	$\hat{\beta}_{[n/2]}$	$\hat{\beta}_{m_1}$	$\hat{\beta}_{m_2}$	$\hat{\beta}_{m_3}$	$\hat{\beta}_{[n/2]}$	$\hat{\beta}_{m_1}$	$\hat{\beta}_{m_2}$	$\hat{\beta}_{m_3}$	$\hat{\beta}_{[n/2]}$
64	0.0	0.49	0.51	0.53	0.63	0.24	0.26	0.27	0.37	0.04	0.04	0.04	0.07
64	0.2	0.61	0.63	0.64	0.71	0.32	0.33	0.35	0.43	0.06	0.07	0.07	0.10
64	0.4	0.77	0.78	0.78	0.81	0.43	0.45	0.45	0.52	0.10	0.11	0.11	0.14
128	0.0	0.41	0.42	0.44	0.49	0.16	0.17	0.19	0.28	0.01	0.02	0.02	0.03
128	0.2	0.53	0.55	0.57	0.66	0.23	0.25	0.26	0.34	0.03	0.03	0.04	0.05
128	0.4	0.74	0.74	0.75	0.79	0.35	0.36	0.38	0.44	0.06	0.06	0.07	0.08

Table 10
MSE for FDLS and OLS (Eqs. (9, 10))

n	$d_x = 0.5$					$d_x = 0.8$				$d_x = 1.2$			
	d_e	$\hat{\beta}_{m_1}$	$\hat{\beta}_{m_2}$	$\hat{\beta}_{m_3}$	$\hat{\beta}_{[n/2]}$	$\hat{\beta}_{m_1}$	$\hat{\beta}_{m_2}$	$\hat{\beta}_{m_3}$	$\hat{\beta}_{[n/2]}$	$\hat{\beta}_{m_1}$	$\hat{\beta}_{m_2}$	$\hat{\beta}_{m_3}$	$\hat{\beta}_{[n/2]}$
64	0.0	0.28	0.29	0.30	0.41	0.08	0.09	0.09	0.15	0.01	0.01	0.01	0.01
64	0.2	0.41	0.43	0.44	0.51	0.13	0.14	0.14	0.21	0.01	0.01	0.01	0.02
64	0.4	0.66	0.65	0.65	0.67	0.23	0.23	0.24	0.29	0.02	0.02	0.02	0.03
128	0.0	0.19	0.20	0.21	0.34	0.04	0.04	0.04	0.09	0.00	0.00	0.00	0.00
128	0.2	0.32	0.33	0.34	0.45	0.07	0.08	0.08	0.13	0.00	0.00	0.00	0.01
128	0.4	0.60	0.59	0.59	0.63	0.15	0.15	0.16	0.21	0.01	0.01	0.01	0.01

a highly challenging task when integration orders are unknown, and in any case preliminary estimates of β will still be needed and the capacity of FDLS to improve on OLS here is still an advantage. Certainly, while biases still remain high in many cases under FDLS, our experiment shows that we can do better with FDLS than OLS, in terms not only of bias but MSE for a range of bandwidths.

In a previous, extended, version of the paper we included also Monte Carlo information (available upon request) on the procedures of Section 4 for testing integration orders of raw data. In short, *LR*-type procedures enjoy some superiority in terms of empirical size, as is consistent with the existing literature on higher order asymptotics for these tests in other settings.

Finally we consider the performance of the Hausman test. One thousand series of lengths $n = 64, 128, 256$ were generated according to (32/33); the null of no-cointegration is identified by $d_e = d_x$. The test is based upon comparison of $\hat{\hat{\delta}}_x$ and $\hat{\hat{\delta}}_*$, where both estimates are evaluated on differenced data; power and size refer to 5% critical values. We considered $d_x = 0.8, 1.2$ and $d_e = 0.0, 0.2, 0.4, d_x$, and also the $I(1)/I(0)$ case $d_x = 1$, $d_e = 0, 1$. Empirical size and power are reported in Table 11.

Table 11
Size and power of the Hausman test at 5% (Eq. (30))

d_e	d_x	$n=64$			$n=128$			$n=256$		
		10 (%)	20 (%)	30 (%)	20 (%)	30 (%)	40 (%)	30 (%)	40 (%)	50 (%)
0.0	0.8	30.3	33.0	35.2	44.3	51.4	56.5	64.5	75.1	80.0
0.2	0.8	25.4	25.1	29.1	32.3	40.9	42.7	51.3	59.3	64.2
0.4	0.8	24.1	18.0	19.4	22.1	27.2	29.5	33.6	38.0	45.4
0.8	0.8	22.9	15.1	14.3	12.9	13.2	10.1	12.9	9.9	9.7
0.0	1	38.4	46.8	49.2	65.5	74.5	81.1	87.4	92.7	94.2
1	1	21.4	14.6	11.1	14.0	12.8	10.5	11.3	11.6	7.8
0.0	1.2	55.6	64.2	62.1	83.7	90.4	94.2	97.9	99.3	99.8
0.2	1.2	47.2	56.0	57.9	76.9	83.9	88.6	95.1	97.7	98.9
0.4	1.2	39.7	46.1	47.2	65.1	71.5	76.9	87.9	92.1	96.2
1.2	1.2	19.4	15.3	11.7	15.0	13.7	10.8	12.3	11.6	10.4

Sizes are larger than one might like, but not unsatisfactory, and they improve with increasing n, as do the powers, which also noticeably increase with $d_x - d_e$; for example, in the $I(1)/I(0)$ case with $n=128$ and $s=40$, the size under the null is slightly above 10% whereas the power exceeds 80%.

Acknowledgements

A preliminary version of the empirical analysis of Section 5 appears in the unpublished paper Robinson and Marinucci (1997). We are grateful to an editor and two referees for comments that greatly improved the presentation. The first author's research was supported by MURST; the second author's research was supported by a Leverhulme Trust Personal Research professorship and by ESRC grant R000235892.

References

Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models. Journal of Political Economy 95, 1062–1089.

Chan, N.H., Terrin, N., 1995. Inference for unstable long-memory processes with applications to fractional unit root autoregressions. Annals of Statistics 23, 1662–1683.

Engle, R., 1974. Band spectrum regression. International Economic Review 15, 1–11.

Engle, R., Granger, C., 1987. Cointegration and error correction: representations, estimation and testing, Econometrica 55, 251–276.

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 models. Journal of Time Series Analysis 4, 221–238.

Hannan, E.J., 1963. Regression for time series with errors of measurement. Biometrika 50, 293–302.

Hausman, J., 1978. Misspecification tests in econometrics. Econometrica 46, 1251–1271.

Hurvich, C.M., Deo, R., Brodsky, J., 1998. The mean squared error of Geweke and Porter-Hudak's estimates of the memory parameter of a long memory time series. Journal of Time Series Analysis 19, 19–46.

Johansen, S., 1991. Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. Econometrica 59, 1551–1580.

Künsch, H.R., 1987. Statistical aspects of self-similar processes. Proceedings of the First World Congress of the Bernoulli Society, VNU Science Press, Vol. 1, pp. 67–74.

Lobato, I.G., 1996. Multivariate analysis of long memory in the frequency domain, Ph.D. Thesis, University of London.

Lobato, I.G., 1999. A semiparametric two-step estimation on a multivariate long memory model. Journal of Econometrics 90, 129–153.

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

Marinucci, D., Robinson, P.M., 1999a. Alternative forms of fractional brownian motion. Journal of Statistical Planning and Inference 80, 111–122.

Marinucci, D., Robinson, P.M., 1999b. Finite sample improvements in statistical inference with $I(1)$ processes. Journal of Applied Econometrics, forthcoming.

Marinucci, D., Robinson, P.M., 2000. Weak convergence of multivariate fractional processes. Stochastic Processes and their Applications 86, 103–120.

Phillips, P.C.B., 1991a. Optimal inference in cointegrated systems. Econometrica 59, 283–306.

Phillips, P.C.B., 1991b. Spectral regression for cointegrated time series. In: Barnett, W.A., Powell, J., Tauchen, G., (Eds.), Nonparametric and Semiparametric Methods in Econometrics and Statistics. Cambridge University Press, Cambridge, pp. 413–435.

Phillips, P.C.B., Ouliaris, S., 1988. Testing for cointegration using principal components methods. Journal of Economics Dynamics and Control 12, 205–230.

Phillips, P.C.B., Hansen, B.E., 1990. Statistical inference in instrumental variable regression with $I(1)$ variables. Review of Economic Studies 57, 99–125.

Robinson, P.M., 1973. Stochastic difference equations with non-integral differences. Advances in Applied Probability 6, 524–545.

Robinson, P.M., 1988. The stochastic difference between econometric statistics. Econometrica 56, 531–548.

Robinson, P.M., 1994a. Semiparametric analysis of long-memory time series. Annals of Statistics 22, 515–539.

Robinson, P.M., 1994b. Efficient 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.

Robinson, P.M., 1998. Comment on real and spurious long memory properties of stock market data, by I.G. Lobato and N.E. Savin. Journal of Business and Economic Statistics 16, 261–282.

Robinson, P.M., Marinucci, D., 1997. Semiparametric frequency-domain analysis of fractional cointegration, preprint.

Robinson, P.M., Marinucci, D., 1999. Narrow-band analysis of nonstationary processes, preprint.

Robinson, P.M., Marinucci, D., 2000. The averaged periodogram for nonstationary vector time series. Statistical Inference for Stochastic Processes 3, 149–160.

Robinson, P.M., Yajima, Y., 2000. Determination of cointegrating rank in fractional systems. Journal of Econometrics, forthcoming.

Stock, J.H., 1987. Asymptotic properties of least squares estimators of cointegrating vectors. Econometrica 55, 1035–1056.

Velasco, C., 1999a. Non-stationary log-periodogram regression. Journal of Econometrics 91, 325–371.

Velasco, C., 1999b. Gaussian semiparametric estimation of non-stationary time series. Journal of Time Series Analysis 20, 87–127.