DEPARTMENT OF ECONOMICS

Working Paper

Optimal Residual Based Tests for Fractional Cointegration and Exchange Rate Dynamics

Morten Ørregaard Nielsen

Working Paper No. 2002-7

ISSN 1396-2426

UNIVERSITY OF AARHUS • DENMARK

INSTITUT FOR ØKONOMI

AFDELING FOR NATIONALØKONOMI - AARHUS UNIVERSITET - BYGNING 350
8000 AARHUS C - ☎ 89 42 11 33 - TELEFAX 86 13 63 34

WORKING PAPER

Optimal Residual Based Tests for Fractional Cointegration
and Exchange Rate Dynamics

Morten Ørregaard Nielsen

Working Paper No. 2002-7

DEPARTMENT OF ECONOMICS

SCHOOL OF ECONOMICS AND MANAGEMENT - UNIVERSITY OF AARHUS - BUILDING 350
8000 AARHUS C - DENMARK ☎ +45 89 42 11 33 - TELEFAX +45 86 13 63 34

Optimal Residual Based Tests for Fractional Cointegration and Exchange Rate Dynamics

Morten Ørregaard Nielsen
Department of Economics
Building 322, University of Aarhus
DK-8000 Aarhus C
Denmark
E-mail: monielsen@econ.au.dk

September 10, 2002

Abstract

We propose a Lagrange Multiplier test of the null hypothesis of cointegration in fractionally cointegrated models. The test statistic utilizes fully modified residuals to cancel the endogeneity and serial correlation biases, and we show that standard asymptotic properties apply under the null and under local alternatives. With *i.i.d.* Gaussian errors the asymptotic Gaussian power envelope of all (unbiased) tests is achieved by the one-sided (two-sided) test. The finite sample properties are illustrated by a Monte Carlo study. In an application to the dynamics among exchange rates for seven major currencies against the US dollar, mixed evidence of the existence of a cointegrating relation is found.

JEL Classification: C12; C22; C32

Keywords: Cointegration Test; Fully Modified Estimation; Nonstationarity; Optimal Test; Power Envelope

1 INTRODUCTION

In this paper we propose a Lagrange Multiplier (LM) test of the null hypothesis of cointegration in fractionally cointegrated models. In nonstationary and possibly cointegrated models, estimators and test statistics are often found to have nonstandard distributional properties when the null is nested in the autoregressive alternatives typically considered in the literature. In contrast, we show that by embedding the model of interest in a general $I(d)$ framework, the LM test statistic regains the standard distributional properties and uniform optimality properties well known from simpler models.

The analysis of cointegration has been a very active area of research in the econometrics and time series literature in the last 20 years, starting with the seminal contributions by Granger (1981) and Engle & Granger (1987). Most of this work has considered the $I(1) - I(0)$ type of cointegration in which linear combinations of two or more $I(1)$ variables are $I(0)$. A process is labelled $I(0)$ if it is covariance stationary and has spectral density that is bounded and bounded away from zero at the origin, and $I(1)$ if the first differenced series is $I(0)$. If y_t and x_t are $I(1)$, and hence in particular nonstationary (unit root) processes, but there exists a process e_t which is $I(0)$ and a fixed β such that

$$y_t = \beta' x_t + e_t, \tag{1}$$

then y_t and x_t are said to be cointegrated. Thus, the nonstationary series move together in the sense that a linear combination of them is stationary and a common stochastic trend is shared. Testing for cointegration in this framework amounts to testing stationarity of the unobserved residual process e_t against a unit root alternative, see e.g. Shin (1994), Jansson (2001), and the references therein.

The above notion of cointegration is based on the knife-edge distinction between $I(1)$ and $I(0)$ processes. However, many economic and financial time series exhibit strong persistence without exactly possessing unit roots, for some recent evidence see e.g. Diebold & Rudebusch (1989), Baillie & Bollerslev (1994), Baillie (1996), Lobato & Velasco (2000), and Marinucci & Robinson (2001). This has led to the consideration of the class of fractionally integrated processes, which is more general than $I(1)$ and still admits a criterion for linear co-movement of series. Thus, a process is fractionally integrated of order d, denoted $I(d)$, if its d'th difference is $I(0)$. Here, d may be any real number, i.e. $d = 0$ or $d = 1$ are special cases. For a precise statement, x_t is $I(d)$ if

$$\Delta^d x_t = u_t \mathbb{I}(t \geq 1) = u_t^{\#}, \tag{2}$$

or equivalently, inverting (2),

$$x_t = \Delta^{-d} u_t^{\#}, \tag{3}$$

defining $u_t^{\#} = u_t \mathbb{I}(t \geq 1)$, where u_t is $I(0)$, $\mathbb{I}(\cdot)$ denotes the indicator function, and the fractional difference operator $\Delta^d = (1-L)^d$ is defined by its binomial expansion

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} L^j, \qquad \Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt, \tag{4}$$

in the lag operator L ($Lx_t = x_{t-1}$). With the definition (2) or (3), x_t is a type II fractionally integrated process, which is nonstationary for all d but asymptotically stationary for $d < 1/2$, see Marinucci & Robinson (1999). Following the original idea by Granger (1981), a natural generalization of the cointegration concept is to assume that the raw series are $I(d)$ and that a certain linear combination is $I(d-b)$, with $d \geq b$ positive real numbers. This is denoted $CI(d,b)$.

To fix ideas, consider the simple system

$$\Delta^{d-b+\theta}\left(y_{1t} - \beta' y_{2t}\right) = u_{1t}^{\#}, \tag{5}$$

$$\Delta^{d} y_{2t} = u_{2t}^{\#}, \tag{6}$$

where $u_t = (u_{1t}, u_{2t}')'$ is $I(0)$. In this model y_t is $CI(d, b-\theta)$ and the cointegration vector is given by $(1, -\beta')$. Clearly, this allows the study of co-movement among persistent series much more generally than in the standard unit root based $I(1) - I(0)$ cointegration framework. In the present paper, we assume that d and b are known a priori and satisfy $d \geq b > 3/4$.

We wish to test the hypothesis $H_0 : \theta = 0$, i.e. setting $d = b = 1$ can be seen as an alternative to testing for stationarity of the residuals in (1). If the null hypothesis is changed slightly in this setup, the properties of the process y_t do not change as dramatically as in the standard cointegration model in which the relation (1) is either perfectly cointegrating, i.e. $CI(1,1)$, or spurious. A notion of near-cointegration does exist in the unit root based $I(1) - I(0)$ cointegration literature, which offers some smoothing of the gap between $CI(1,1)$ and spurious regression, e.g. Jansson & Haldrup (2001). However, the test statistics in that framework still have nonstandard distributional properties.

We show that in our fractional integration framework much more desirable properties obtain than in (1). Our test can be considered an extension of the univariate LM tests in Robinson (1991, 1994), Agiakloglou & Newbold (1994), and Tanaka (1999), among others, who considered testing for a unit root in a fractional integration framework, i.e. testing on the parameter d in (2) in the frequency and time domains. They showed that their tests have standard asymptotic distributions and, under Gaussianity, that their tests enjoy optimality properties. Simulations in Tanaka (1999) showed that, in finite samples, the time domain tests are superior to Robinson's (1994) frequency domain LM test with respect to both size and power.

Undoubtedly there exist Wald and likelihood ratio versions of our LM test, which have the same asymptotic properties as our test even though their finite sample properties may differ. However, we consider only the time domain LM test for fractional cointegration with the usual computational motivation that the model only needs to be estimated under the null hypothesis. As we shall see below, in the important special case $d = b = 1$ the computation of the LM test statistic does not require any fractional differencing, and indeed all that is needed in this case are the residuals from a fully modified regression which can be obtained from readily available computer software.

We show that the likelihood theory in the time domain is tractable and that the ML estimator of the cointegrating vector β, which is required to compute the test statistic, reduces to a version of the fully modified least squares estimator of Phillips & Hansen (1990) and Phillips (1991), see also Kim & Phillips (2001) for a fractional cointegration version. We then show that the LM test can be calculated using the residuals from the fully modified regression and establish the desirable distributional properties and optimality properties of the test. In particular, the test statistic is consistent and asymptotically normal or chi-squared distributed, and under the additional assumption of Gaussianity the test is locally most powerful. Indeed, we show that in the special case with *i.i.d.* Gaussian errors, the asymptotic Gaussian power envelope of all (unbiased) tests is achieved by the one-sided (two-sided) version of our test, i.e. the one-sided (two-sided) test is uniformly most powerful among all (unbiased) tests. In a simulation study we find that the finite sample rejection frequencies are reasonable but well below the asymptotic local power for samples of size $n = 200$, and much closer to the asymptotic local power for $n = 500$.

Our new methodology is applied to the analysis of exchange rate dynamics following Baillie & Bollerslev (1989, 1994). Previous studies have focused on the estimation of the cointegration vector and the memory parameter of the equilibrium errors, but no formal testing of the hypothesis of fractional cointegration has been done. We concentrate on testing for the presence of (fractional) cointegration with various specifications of d and b. Our findings are not decisive, but we do find some evidence of cointegration among a system of exchange rates for seven major currencies against the US Dollar. In particular, we cannot reject (against fractional alternatives) that the exchange rates can be described by a standard $I(1) - I(0)$ cointegration model when the errors (i.e. u_{1t} and u_{2t} in (5) and (6) above) are allowed to follow autoregressive processes of order one.

The remainder of the paper is laid out as follows. Section 2 sets up the model of fractional cointegration. In section 3 we consider the estimation of the cointegrating vector, derive the LM test statistic, and establish the desirable distributional properties. In section 4 we derive the asymptotic Gaussian power envelopes for the one-sided and two-sided testing problems and show that they coincide with the local asymptotic power functions of the one-sided and two-sided LM tests. Section 5 presents the results of the Monte Carlo study and in section 6 we provide the empirical application to exchange rate dynamics. Section 7 offers some concluding remarks. All proofs are collected in the appendix.

2 A MODEL OF FRACTIONAL COINTEGRATION

Suppose we observe the K-vector time series $\{y_t, t = 1, 2, ..., n\}$, which we partition as y_{1t} (scalar) and y_{2t} ($(K-1)$-vector). We consider a triangular model of fractional cointegration in the spirit of the Phillips (1991) triangular system. Thus, let y_t be generated by the fractionally cointegrated system

$$y_{1t} = \beta' y_{2t} + z_t, \quad t = 1, 2, ..., \tag{7}$$

$$\Delta^{d-b+\theta} z_t = u_{1t}^{\#}, \quad t = 1, 2, ..., \tag{8}$$

$$\Delta^{d} y_{2t} = u_{2t}^{\#}, , \quad t = 1, 2, ..., \tag{9}$$

where z_t is the (unobserved) deviation from the cointegrating relation and $u_t = (u_{1t}, u'_{2t})'$ is an error component. We allow the error components u_{1t} and u_{2t} to be contemporaneously correlated and possibly weakly dependent, c.f. Assumption 1 below.

The system (7) $-$ (9) generalizes the standard triangular cointegration model. The series share fractionally integrated stochastic trends of orders $I(d)$ and $I(d-b)$, and the linear combination $(1, -\beta')$ eliminates the most persistent one. Equation (7) can be regarded as an equilibrium relationship between the $I(d)$ components of y_t. Under the null, $\theta = 0$, the deviations from equilibrium constitute an $I(d-b)$ process, and when $d = b$ the deviations are only weakly dependent, so this is a case of special interest. The model could be extended to multidimensional cointegrating relationships as in Jeganathan (1999), where the estimation of the cointegration rank and cointegrating vectors are of interest. However, most empirical studies consider a single cointegrating relation among two or more variables, e.g. Cheung & Lai (1993), Baillie & Bollerslev (1994), Dueker & Startz (1998), Marinucci & Robinson (2001), and Kim & Phillips (2001). Thus, we consider only the case of a single cointegrating relationship in this paper to keep focus on optimal testing of hypotheses on θ.

The model is assumed to satisfy the following assumption on the error process.

Assumption 1 *We consider four typical specifications for the error component u_t. In each case, the innovations $e_t = (e_{1t}, e'_{2t})' \sim i.i.d.(0, \Sigma)$ with finite fourth moment and Σ is a positive definite matrix which we partition conformably as*

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma'_{21} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}. \tag{10}$$

0. $u_t \sim i.i.d.$ or equivalently $u_t = e_t$.

1. *u_{1t} follows the stationary AR(p) process*

$$g(L)\, u_{1t} = e_{1t}, \quad t = 1, 2, ..., \tag{11}$$

and $u_{2t} = e_{2t}$.

2. *$u_{1t} = e_{1t}$, and u_{2t} follows the $(K-1)$-dimensional stationary VAR(p) process*

$$G(L)\, u_{2t} = e_{2t}, \quad t = 1, 2, ... \tag{12}$$

3. *u_t follows the K-dimensional block diagonal stationary VAR(p) process*

$$g(L)\, u_{1t} = e_{1t}, \quad t = 1, 2, ..., \tag{13}$$

$$G(L)\, u_{2t} = e_{2t}, \quad t = 1, 2, ... \tag{14}$$

In cases 2-4, $g(z)$ and $G(z)$ are lag polynomials of order p with coefficients gathered in γ_1 and γ_2, respectively, and $G(1)$ has full rank (no cointegration among the components of y_{2t}).

In the following we write $A(z) = \text{diag}(g(z), G(z))$ as shorthand for the lag polynomial in Assumption 1.3. It would be straightforward to extend Assumption 1.3 to $A(L)\, u_t = e_t$, for a general lag polynomial $A(z)$ of order p, where $A(1)$ has full rank. Applying the formulae in Hosking (1980), the results in Lemma 1 and the following theorems could be extended to cover this more general case. However, the structure imposed by Assumption 1.3 seems relevant and its interpretation is natural.

In our model the constants d and b are prespecified. In particular, we assume that $d \geq b > 3/4$ such that the series are nonstationary and cointegration reduces the integration order by more than 3/4. Assuming that b is known a priori is natural as it effectively specifies the null for our LM test and thus, according to the LM principle, there is no need to estimate b. If d is not known a priori it can be estimated in a preliminary step as in, e.g., Cheung & Lai (1993), Baillie & Bollerslev (1994), Marinucci & Robinson (2001), and Kim & Phillips (2001). Efficient procedures have been developed to estimate d in fractionally integrated time series models, e.g. Sowell (1992) (exact ML) and Tanaka (1999) (conditional ML).

Our objective is to test the hypothesis

$$H_0 : \theta = 0 \tag{15}$$

against $H_1 : \theta > 0$ or $H_2 : \theta \neq 0$ in the model $(7) - (9)$. In particular, $d = b = 1$ generates a standard $I(1) - I(0)$ cointegrated system under the null, so this is a test of the null of cointegration in the usual sense, but the fractional alternatives against which the test is directed are new. Thus, a test of (15) can be considered an alternative to testing stationarity of the residuals in (1), which has been standard in the literature, see e.g. Shin (1994), Jansson (2001), and the references therein. Another important case, for $d \geq 1.25$ and some small user-chosen $\varepsilon > 0$, is the one-sided test of (15) with $b = d - 1/2 + \varepsilon$, i.e. $d - b = 1/2 - \varepsilon$, which is a test for the existence of an (asymptotically) stationary cointegrating relation against the alternative that no stationary cointegrating relation exists (though a nonstationary but mean-reverting cointegrating relation with $1/2 \leq d - b < 1$ may still exist). Finally, for $d \geq 1$ and some small user-chosen $\varepsilon > 0$, it is of interest to conduct a one-sided test of (15) with $b = d - 1/4 + \varepsilon$, i.e. $d - b = 1/4 - \varepsilon$, as a border case for square integrability of the spectral density of the equilibrium errors and asymptotic normality of the autocovariances of the equilibrium errors, see e.g. Fox & Taqqu (1986).

Choosing $d = b = 1$ also suggests applying a test of (15) as a valuable diagnostics tool in a standard $I(1) - I(0)$ cointegration analysis. In this context, rejecting (15) should be taken either as evidence of a drastically misspecified dynamic structure or as a suggestion to employ an actual fractional cointegration analysis. Thus, the test could be thought of as a general test for misspecification of the model. If, for example, y_{1t} and y_{2t} are related by some complicated nonlinear filter and a linear model is imposed, then it is plausible that long-range dependence could be introduced in the residuals as a result of this misspecification.

3 TESTING FRACTIONAL COINTEGRATION

The log-likelihood function of the model $(7)-(9)$ under Assumption 1.3 (the most general case) and Gaussianity of the errors is

$$L(\theta, \beta, \Sigma, \gamma) = -\frac{n}{2}\ln|\Sigma| - \frac{1}{2}\sum_{t=1}^{n}\begin{pmatrix} g(L)\Delta^{d-b+\theta}z_t \\ G(L)\Delta^d y_{2t}\end{pmatrix}'\Sigma^{-1}\begin{pmatrix} g(L)\Delta^{d-b+\theta}z_t \\ G(L)\Delta^d y_{2t}\end{pmatrix} \tag{16}$$

bearing in mind the truncation in our definition of fractionally integrated processes, e.g. $G(L)\Delta^d y_{2t} = e_{2t}^{\#}$ by (9) and (14). The log-likelihood in (16) is equal to the sum of the marginal log-likelihood

$$-\frac{n}{2}\ln|\Sigma_{22}| - \frac{1}{2}\sum_{t=1}^{n}G(L)\Delta^d y_{2t}'\Sigma_{22}^{-1}G(L)\Delta^d y_{2t} \tag{17}$$

and the conditional log-likelihood

$$-\frac{n}{2}\ln\sigma_{1.2}^2 - \frac{1}{2\sigma_{1.2}^2}\sum_{t=1}^{n}\left(g(L)\Delta^{d-b+\theta}\left(y_{1t} - \beta' y_{2t}\right) - \sigma_{21}'\Sigma_{22}^{-1}G(L)\Delta^d y_{2t}\right)^2, \tag{18}$$

where $\sigma_{1.2}^2 = \sigma_{11}^2 - \sigma_{21}' \Sigma_{22}^{-1} \sigma_{21}$ is the variance of $e_{1.2t} = e_{1t} - \sigma_{21}' \Sigma_{22}^{-1} e_{2t}$, which is e_{1t} centered about its mean conditional on e_{2t}. The asymptotic results derived later impose only Assumption 1 on the error process. Gaussianity is not necessary for most of our results and is used only to choose a likelihood function and to derive optimality properties.

From the conditional likelihood (18), the MLE of β under the null, $\theta = 0$, is recognized to be the NLS estimator in the augmented regression

$$\Delta^{d-b} y_{1t} = \beta' \Delta^{d-b} y_{2t} + (g(L) - 1) \Delta^{d-b} \left(y_{1t} - \beta' y_{2t} \right) + c' G(L) \Delta^d y_{2t} + e_{1.2t}, \tag{19}$$

see Phillips & Loretan (1991) for a discussion of the equivalent estimator in the standard $I(1) - I(0)$ cointegration framework. Presumably, the lagged equilibrium errors in (19) could be replaced by leaded $\Delta^d y_{2t}$, as demonstrated by Saikkonen (1991) in the standard cointegration framework, and the resulting regression could be estimated by OLS.

Under Assumption 1.2 where $g(z) = 1$, i.e. when there is no autoregressive term in the equilibrium errors, the estimation of (19) reduces to OLS on

$$\Delta^{d-b} y_{1t} = \beta' \Delta^{d-b} y_{2t} + \sum_{k=0}^{p} c_k \Delta^d y_{2t-k} + e_{1.2t}. \tag{20}$$

This simplification is even stronger under Assumption 1.0 where $p = 0$ in (20) and the lagged fractionally differenced y_{2t} disappear. The simplification (20) is especially useful in many applications where cointegration is a result of rational expectations theory, i.e. that deviations from equilibrium in time t should be unpredictable based on information up to time $t-1$, which in our framework implies $d = b$ and $g(z) = 1$.

Equivalently, (20) is OLS in the (infeasible) regression

$$\Delta^{d-b} y_{1t}^* = \beta' \Delta^{d-b} y_{2t} + e_{1.2t}, \tag{21}$$

where $y_{1t}^* = y_{1t} - \sigma_{21}' \Sigma_{22}^{-1} \sum_{k=0}^{p} \Delta^b y_{2t-k}$. This is the fully modified least squares method of Phillips & Hansen (1990) and Phillips (1991), which was developed for fractional cointegration by Kim & Phillips (2001). In contrast to our restrictions on d and b, Kim & Phillips (2001) require $2d - b > 1, d \geq 1$ in their fully modified method and further that $b \geq 1$ in the likelihood analysis of their model. Thus, Kim & Phillips (2001) limit the strength of the cointegrating relation by bounding $b < 2d - 1$ from above, and in particular they exclude the $CI(1,1)$ case. We assume at least $b > 3/4$ in our analysis, since our estimation problem under the null has been transformed into a regression between $I(b)$ processes with $I(0)$ errors, (19) – (20). Thus the necessity of at least $b > 1/2$ becomes clear, since otherwise the estimator of β becomes inconsistent as demonstrated by e.g. Marinucci & Robinson (2001, p. 231).

Note that if OLS is applied to (7) directly, which has often been the case in the literature, see e.g. Cheung & Lai (1993) or Baillie & Bollerslev (1994), it introduces a bias unless $\sigma_{21} = 0$ and $g(z) = 1$. Indeed, if $\sigma_{21} = 0$ and $g(z) = 1$ hold, y_{2t} is strictly exogenous and inference on θ (and estimation of the parameter β) will depend only on the part of the likelihood attributed to (7). In particular, the MLE of β reduces to OLS on (7) and we can apply the univariate methods of Robinson (1994) and Tanaka (1999). This is not the case when $\sigma_{21} \neq 0$ or $g(z) \neq 1$ because of the well known endogeneity and serial correlation biases, see e.g. Phillips (1991).

Returning to the full model, the normalized score statistic is found by differentiating (16) or (18) with respect to θ and evaluating the resulting expression under the null,

$$\begin{aligned} S_n &= \frac{1}{\sqrt{n}} \frac{\partial L(\theta, \beta, \Sigma, \gamma)}{\partial \theta}\bigg|_{\theta=0,\beta=\hat{\beta},\Sigma=\hat{\Sigma},\gamma=\hat{\gamma}} \\ &= \frac{-1}{\sqrt{n}\hat{\sigma}_{1.2}^2} \sum_{t=1}^{n} \left(\ln(\Delta) \left(\hat{g}(L) \Delta^{d-b} (y_{1t} - \hat{\beta}' y_{2t}) \right) \right) \left(\hat{g}(L) \Delta^{d-b} (y_{1t} - \hat{\beta}' y_{2t}) - \hat{c}' \hat{G}(L) \Delta^d y_{2t} \right) \end{aligned} \tag{22}$$

where $\hat{g}(z)$ and $\hat{G}(z)$ are evaluated at $\hat{\gamma}_1$ and $\hat{\gamma}_2$, respectively. Using that $\ln(1-z) = -\sum_{j=1}^{\infty} j^{-1} z^j$ and defining the fully modified residuals under $\theta = 0$ as

$$\hat{e}_{1.2t} = \hat{g}(L) \Delta^{d-b} (y_{1t} - \hat{\beta}' y_{2t}) - \hat{c}' \hat{G}(L) \Delta^d y_{2t} \tag{23}$$

and

$$\hat{e}_{1t} = \hat{g}(L) \Delta^{d-b} (y_{1t} - \hat{\beta}' y_{2t}), \tag{24}$$

the score can be written more compactly as

$$\begin{aligned} S_n &= \frac{1}{\sqrt{n}\hat{\sigma}_{1.2}^2} \sum_{t=1}^{n} \sum_{j=1}^{t-1} j^{-1} \hat{e}_{1t-j} \hat{e}_{1.2t} \\ &= \frac{\sqrt{n}}{\hat{\sigma}_{1.2}^2} \sum_{j=1}^{n-1} j^{-1} \left(\hat{C}_{11}(j) - \hat{c}' \hat{C}_{21}(j) \right) \\ &= \sqrt{n} \sum_{j=1}^{n-1} j^{-1} e_1' \hat{\Sigma}^{-1} \hat{C}(j) e_1, \end{aligned} \tag{25}$$

where $\hat{C}_{ab}(j) = n^{-1} \sum_{t=j+1}^{n} \hat{e}_{at} \hat{e}_{bt-j}'$ is the estimated sample autocovariance function, $e_1 = (1, 0')'$ is the selection vector, and we used that $\sigma_{1.2}^{-2} = \Sigma^{11}$ and $-\sigma_{1.2}^{-2} \sigma_{21}' \Sigma_{22}^{-1} = \Sigma^{12}$, where Σ^{ab} is the (a, b)'th block of Σ^{-1} for $a, b = 1, 2$.

The asymptotic distribution of the score statistic S_n under the null (15) is considered next.

Theorem 3.1 *Suppose $d \geq b > 3/4$ in the model (7) – (9) and let S_n be defined by (25). Under $H_0 : \theta = 0$ and Assumption 1.0,*

$$S_n \xrightarrow{D} N\left(0, \frac{\pi^2}{6} \frac{\sigma_{11}^2}{\sigma_{1.2}^2}\right). \tag{26}$$

Under $H_0 : \theta = 0$ and Assumption 1.i, S_n is asymptotically Gaussian with mean zero and variance

$$\frac{\pi^2}{6}\frac{\sigma_{11}^2}{\sigma_{1.2}^2} - \operatorname{vec}\left(\Sigma^{-1}e_1e_1'\Sigma\Phi_{i1}', ., \Sigma^{-1}e_1e_1'\Sigma\Phi_{ip}'\right)' H_i \times \left(H_i'\left(\Gamma_i \otimes \Sigma^{-1}\right)H_i\right)^{-1} H_i' \operatorname{vec}\left(\Sigma^{-1}e_1e_1'\Sigma\Phi_{i1}', ., \Sigma^{-1}e_1e_1'\Sigma\Phi_{ip}'\right) \tag{27}$$

for $i = 1, 2, 3$. Here, Γ_i is the covariance matrix of $(u_t', ..., u_{t-p+1}')'$, $\Phi_{il} = \sum_{j=l}^{\infty} j^{-1}\Psi_{i,j-l}$, $\Psi_{i,k}$ is the k'th term in the Wold representation of u_t normalized such that $\Psi_{i,0} = I_K$, and $H_i = \left(\partial a_1'/\partial\gamma_i, ..., \partial a_p'/\partial\gamma_i\right)'$, where $a_j = \operatorname{vec} A_j$ are the coefficients in the autoregressive representation $A(L)u_t = e_t$.

In the simple bivariate VAR(1) example also considered in the appendix, the variance equations (27) reduce to

$$\frac{\pi^2}{6}\frac{\sigma_{11}^2}{\sigma_{1.2}^2} - \operatorname{vec}\left(\Sigma^{-1}e_1e_1'\Sigma\Phi_{i1}'\right)' H_i \left(H_i'\left(\Sigma^{-1} \otimes \Gamma_i\right)H_i\right)^{-1} H_i' \operatorname{vec}\left(\Sigma^{-1}e_1e_1'\Sigma\Phi_{i1}'\right), \ i = 1, 2, 3,$$

where $\Gamma_i = E\left(u_tu_t'\right)$ can be estimated by $n^{-1}\sum_{t=1}^{n}\hat{u}_t\hat{u}_t'$ and the particular Φ_{i1} and H_i for this example are given in the appendix.

The Fisher information for θ, which is derived in the next theorem, illustrates the standard nature of our testing problem.

Theorem 3.2 *Let the assumptions of Theorem 3.1 be satisfied and assume that $\{e_t\}$ is Gaussian. Under Assumption 1.0 the Fisher information for θ is*

$$\mathcal{I}_0 = -\lim_{n\to\infty} E\left(\frac{1}{n}\frac{\partial^2 L(\theta, \beta, \Sigma)}{\partial\theta\partial\theta'}\right) = \frac{\pi^2}{6}\frac{\sigma_{11}^2}{\sigma_{1.2}^2}, \tag{28}$$

and under Assumption 1.i, i=1,2,3, the Fisher information for θ is

$$\mathcal{I}_i = \frac{\pi^2}{6}\frac{\sigma_{11}^2}{\sigma_{1.2}^2} - \operatorname{vec}\left(\Sigma^{-1}e_1e_1'\Sigma\Phi_{i1}', ., \Sigma^{-1}e_1e_1'\Sigma\Phi_{ip}'\right)' H_i \times \left(H_i'\left(\Gamma_i \otimes \Sigma^{-1}\right)H_i\right)^{-1} H_i' \operatorname{vec}\left(\Sigma^{-1}e_1e_1'\Sigma\Phi_{i1}', ., \Sigma^{-1}e_1e_1'\Sigma\Phi_{ip}'\right). \tag{29}$$

To assess the local power properties of the test, we derive the asymptotic distribution under the sequence of local alternatives $\theta_{1n} = \delta/\sqrt{n}$.

Theorem 3.3 *Under the assumptions of Theorem 3.1 and $\theta = \delta/\sqrt{n}$,*

$$S_n \xrightarrow{D} N\left(\delta\mathcal{I}_i, \mathcal{I}_i\right) \tag{30}$$

as $n \to \infty$, where $\mathcal{I}_i$ is defined in Theorem 3.2.

Consider again briefly the special case where it is known that $\sigma_{21} = 0$. In that case the score (25) and the distributions in Theorem 3.3 coincide with the ones obtained by Tanaka (1999). That is, we can apply the test of Tanaka (1999) to the residuals in (24), and Tanaka's (1999) *i.i.d.* result obtains under Assumptions 1.0 and 1.2 and his result for autocorrelated errors obtains under Assumptions 1.1 and 1.3. Thus, when $\sigma_{21} = 0$, our test has the same functional form and distribution as Tanaka's (1999) test, which is based on more information (β known), and therefore our test shares the asymptotic optimality properties of that test when $\sigma_{21} = 0$.

In practice, to construct an approximate size α test of H_0 against $H_1 : \theta > 0$ under Assumption 1.i, we compute the statistic

$$LM_{i1} = \frac{1}{\sqrt{\mathcal{I}_i}} S_n \xrightarrow{D} N\left(\delta\sqrt{\mathcal{I}_i}, 1\right) \tag{31}$$

under $\theta = \delta/\sqrt{n}$ as $n \to \infty$, and compare it to the $100(1-\alpha)\%$ point of the standard normal distribution. To test against the two-sided alternative $H_2 : \theta \neq 0$ under Assumption 1.i, at approximate size α, we compute

$$LM_{i2} = LM_{i1}^2 \xrightarrow{D} \chi_1^2\left(\delta^2 \mathcal{I}_i\right) \tag{32}$$

under $\theta = \delta/\sqrt{n}$ as $n \to \infty$, and compare it to the $100(1-\alpha)\%$ point of the central χ_1^2 distribution.

A useful feature of the asymptotic distributions (31) and (32) is that they are free of the parameters d and b. Since d and b are assumed known a priori, their effect is neutralized by suitable differencing. This shows that simple asymptotic inference about θ can be carried out for any choice of d and b satisfying $d \geq b > 3/4$.

The calculation of the tests may seem to be quite involved as p gets large because of the covariance matrices Φ and Γ. However, for a given parameter value γ we can calculate Φ and Γ (and thus the tests) simply by finding the coefficients in the Wold representation of $\hat{u}_t$, then directly evaluate the sums in Φ, and set Γ equal to the sample covariance matrix of $(\hat{u}_t', ..., \hat{u}_{t-p+1}')'$. Another possibility is to employ the following numerical approximation to the one-sided test,

$$\widehat{LM}_{i1} = -\sqrt{n} \sum_{t=1}^{n} \hat{e}_{1.2t} \frac{\partial \hat{e}_{1t}}{\partial \theta} \Bigg/ \sqrt{\sum_{t=1}^{n} \left(\frac{\partial \hat{e}_{1t}}{\partial \theta}\right)^2 \sum_{t=1}^{n} \hat{e}_{1.2t}^2}\,\Bigg|_{H_0}, \tag{33}$$

which follows by noting that $\partial \hat{e}_{1t}/\partial\theta = -\sum_{j=1}^{t-1} j^{-1} \hat{e}_{1t-j}$ and comparing with (25) and (31).

From Theorem 3.3, we can easily calculate the asymptotic local power functions of the one-sided and two-sided tests (31) – (32). This is stated as a corollary.

Corollary 3.1 *Under the assumptions of Theorem 3.3 it holds that, under* $\theta = \delta/\sqrt{n}$*,*

$$P(LM_{i1} > Z_{1-\alpha}) \rightarrow \Phi\left(Z_\alpha + |\delta| \sqrt{\mathcal{I}_i}\right), \quad (34)$$

$$P\left(LM_{i2} > \chi^2_{1,1-\alpha}\right) \rightarrow 1 - F_{\lambda_i}\left(\chi^2_{1,1-\alpha}\right), \quad (35)$$

where $Z_{1-\alpha}$ *and* $\chi^2_{1,1-\alpha}$ *are the* $100(1-\alpha)\%$ *points of the standard normal and central* χ^2_1 *distributions, respectively, and* Φ *and* F_{λ_i} *are the distribution functions of the standard normal distribution and the noncentral* χ^2_1 *distribution with noncentrality parameter* $\lambda_i = \delta^2 \mathcal{I}_i$*.*

Figure 1 shows asymptotic local power functions for $d = b = 1$ and a variety of first order autoregressive specifications and contemporaneous correlation structures. When the correlation is low (left-hand side panels) only the autoregressive term in the equilibrium error (8) has a significant effect. In fact, if the errors are contemporaneously uncorrelated, i.e. $\sigma_{21} = 0$, the power functions for cases 1.0 and 1.2 coincide and the power functions for cases 1.1 and 1.3 coincide. With highly correlated errors (right-hand side panels) the autocorrelation in $\Delta^d y_{2t}$ spills over and has some effect on the power function, though still not as much as the autoregressive term in the equilibrium error. This is well known from standard cointegration analysis. Since the regressors y_{2t} are already heavily trended it makes little difference if the innovations to the stochastic trend are weakly autocorrelated.

Figure 1 about here

It follows from Corollary 3.1 and Theorem 3.2 that the power functions of the tests depend on the covariance matrix of the underlying innovations e_t, such that the power depends on the extent of the endogeneity of the regressors y_{2t}. In particular, under Assumptions 1.0 and 1.1 any correlation between y_{2t} and z_t is exploited by the test to increase power, c.f. equation (28) and Figure 1 (compare the starred lines in the left-hand and right-hand side panels). Note that in case 1.2 the power may increase or decrease with correlated errors. Comparing the solid line in the left-hand and right-hand side panels the power increases when correlation is increased from .6 to .9. However, comparing the starred and solid lines in the upper left-hand side panel shows that in case 1.2 power is decreased with correlation .6 compared to the uncorrelated case. Thus, when correlation is high the first term in $\mathcal{I}_2$, which increases power as correlation increases, dominates the second term, which decreases power due to the spill-over of the autocorrelation via the contemporaneous correlation.

In general the ability of the test to exploit the correlation stands in contrast to the standard $I(1) - I(0)$ framework where the power functions and power envelopes do not depend on Σ, see Jansson & Haldrup (2001) and Jansson (2001). The dependence on Σ is due to the fact that β can be assumed to be known in the derivation of power functions and power envelopes in our fractional setup. That is not the case in the standard $I(1)-I(0)$ framework and thus the standard cointegration tests are unable to exploit the correlation between y_{2t} and z_t to gain power.

As a first optimality result, it follows immediately from Theorems 3.2 and 3.3 that the two-sided test is locally most powerful (LMP) and we state this as a corollary.

Corollary 3.2 *Under the assumptions of Theorem 3.3 and the additional assumption of Gaussianity, the two-sided test statistic* (32) *is locally most powerful in the sense that the noncentrality parameter is maximal.*

Next, we show that much stronger optimality results than the LMP property of Corollary 3.2 can be obtained for the problem of testing (15) when the errors are assumed to be *i.i.d.* Gaussian.

4 ASYMPTOTIC GAUSSIAN POWER ENVELOPES

In this section, we derive the asymptotic Gaussian power envelopes for the one-sided and two-sided testing problems and proceed to show, following Elliott, Rothenberg & Stock (1996) and Tanaka (1999), that the one-sided test is uniformly most powerful (UMP) and, following Nielsen (2001), that the two-sided test is uniformly most powerful unbiased (UMPU).

Assume that the data generating process is (7) – (9) with u_t independent, normally distributed, β and Σ known, and true parameter value $\theta_{0n} = c/\sqrt{n}$ for some fixed $c > 0$. The test of $H_0 : \theta = 0$ against the local alternative $H_1 : \theta_{1n} = \delta/\sqrt{n}$ for some fixed $\delta > 0$ is a test of a simple null against a simple alternative. The Neyman-Pearson Lemma, e.g. Lehmann (1986, chapter 3), states that the test that rejects the null when

$$M_n = n \frac{\sum_{t=1}^n \tilde{u}_{1.2nt}^2 - \sum_{t=1}^n \hat{u}_{1.2nt}^2}{\sum_{t=1}^n \tilde{u}_{1.2nt}^2} \tag{36}$$

becomes large is most powerful. Here, $\tilde{u}_{1.2nt}$ and $\hat{u}_{1.2nt}$ are the residuals (with β and Σ known) under H_0 and H_1 respectively. The next theorem derives the limiting distribution of M_n under local alternatives.

Theorem 4.1 *Let M_n denote the test statistic* (36) *in the model generated by $\theta_{0n} = c/\sqrt{n}$ ($c > 0$ is a fixed scalar). Then, under the sequence of local alternatives $\theta_{1n} = \delta/\sqrt{n}$ ($\delta > 0$ is a fixed scalar), it holds that*

$$M_n \xrightarrow{D} M(c,\delta) = 2\delta\sqrt{\mathcal{I}_0} Z + \delta(2c-\delta)\mathcal{I}_0$$

as $n \to \infty$, where Z is a standard normal variable.

Let the power of M_n be given by $\pi(c,\delta) = P(M(c,\delta) > c_\alpha(\delta))$ under H_{1n} when θ_{0n} is true, where the critical value $c_\alpha(\delta)$ is determined by $P(M(0,\delta) > c_\alpha(\delta)) = \alpha$. Then the power envelope of all one-sided tests is given by $\Pi(\delta) = \pi(\delta,\delta)$, and a test whose power attains the power envelope for all points δ is UMP.

To find a test statistic that applies against two-sided alternatives we invoke the principle of unbiasedness, see Lehmann (1986, chapter 4), to construct a most powerful unbiased test. Unbiasedness requires that the power of the test does not fall below the nominal significance level for any point in the alternative. A test whose power attains the power envelope for all points δ is UMPU.

The following theorem derives the asymptotic Gaussian power envelopes of the one-sided and two-sided testing problems, and shows that these envelopes are achieved by our tests.

Theorem 4.2 *The one-sided asymptotic Gaussian power envelope for all tests of size* α *of* $H_0 : \theta = 0$ *against* $H_1 : \theta_{1n} = \delta/\sqrt{n}$ *(δ a fixed scalar) is given by* (34) *and the two-sided asymptotic Gaussian power envelope for all unbiased tests of size* α *is given by* (35)*. Thus, the one-sided LM test* (31) *is uniformly most powerful (UMP) and the two-sided LM test* (32) *is uniformly most powerful among all unbiased tests (UMPU).*

This result is in stark contrast to the results in the standard $I(1) - I(0)$ cointegration literature. Tests that enjoy optimality properties have been derived in that framework by e.g. Shin (1994) and Jansson (2001) whose tests are LMP and point optimal, respectively, i.e. tests that have maximal power against a single prespecified point in the alternative. However, our criterion is against all possible alternatives.

5 FINITE SAMPLE PERFORMANCE

The local power functions and power envelopes derived above are asymptotic results, and in this section we examine by Monte Carlo experiments whether these asymptotic approximations carry over to finite samples.

The model we have chosen for the simulation study is a bivariate system with $d = b = 1$, i.e.

$$\Delta^\theta (y_{1t} - y_{2t}) = u_{1t}^{\#}, \tag{37}$$

$$\Delta y_{2t} = u_{2t}^{\#}, \tag{38}$$

which is a standard cointegrated model under the null. We consider several specifications for the error process corresponding to each case in Assumption 1 and let e_t be bivariate normal with variances normalized to unity and with contemporaneous correlation 0 or .6. The parameter values for the autoregressive coefficients correspond to those in the upper panels in Figure 1, i.e. $\gamma_1 = .2$ and $\gamma_2 = .5$.

All calculations were made in Ox v3.00 (Doornik (2001)) including the Arfima package v1.01 (Doornik & Ooms (2001)). Throughout, we fix the nominal size (type I error) at .05 and the number of replications at $1,000$. We consider the sample sizes $n = 200$ and $n = 500$. The first is typical for macroeconomic time series, and the latter (or even larger) for financial time series.

We concentrate on comparing the finite sample performance of the one-sided LM test (reported as LM) with the asymptotic local power, but also report results for the size corrected LM test (reported as LMsc). The properties of the estimator of the cointegrating vector β in a similar model were examined by Kim & Phillips (2001), who found that even in samples as small as $n = 100$ the performance of the estimator is very good with respect to bias and variance.

Tables 1-4 show the simulated rejection frequencies of the test statistics (LM and LMsc) for different assumptions on the autocorrelation structure of the errors (as in Figure 1) corresponding to each case outlined in Assumption 1. For comparison, the asymptotic local power, which is equal to the power envelope under Assumption 1.0 by Theorem 4.2, has been calculated from Corollary 3.1 for the same parameter values and is reported under the heading 'Envelope'. The first three columns of each table give the results for contemporaneously uncorrelated errors, whereas in the last three columns the contemporaneous correlation between the errors is .6.

Tables 1-4 about here

First, consider the case where the cointegrating errors u_{1t} are *i.i.d.* and u_{2t} are either *i.i.d.* (Table 1) or follow an AR(1) (Table 3). In these two cases the finite sample rejection frequencies are quite close to the asymptotic local power, even for the small sample size $n = 200$, and especially with contemporaneously uncorrelated errors. In Table 3 the effect of $G(z)$ spills over via the correlation and slightly degrades the size and power compared to the uncorrelated case where there is no spill-over. The insignificance of the specification of u_{2t} is well known from standard cointegration analysis, and is due to the fact that y_{2t} is already highly trended and making the innovations to this $I(d)$ process weakly dependent does not add significantly to this trend.

When u_{1t} is allowed to be autocorrelated as in Tables 2 and 4, where u_{1t} follows an AR(1) process, we know from Corollary 3.1 and Figure 1 that the power of the test degrades and consequently the asymptotic local power functions are much lower than in Tables 1 and 3. The finite sample rejection frequencies reflect this behavior and are well below the asymptotic power for $n = 200$ and also somewhat below the asymptotic power for $n = 500$.

Comparing the middle and right-hand side panels in Tables 1, 2, and 4 shows that the test takes advantage of the correlation between the underlying errors, and the improvement in power when the errors are correlated (right-hand side panels) is evident. The ability of the test to exploit this correlation to increase power even in finite samples is remarkable and contrasts the inability of conventional cointegration tests to exploit this correlation even asymptotically, see Jansson & Haldrup (2001) and Jansson (2001).

In general, the finite sample power functions for samples of size $n = 200$ are reasonable, but well below the asymptotic local power. For samples of size $n = 500$ they are close to the asymptotic local power functions, especially in the absence of an autoregressive term in the equilibrium errors. Thus, one would expect very good performance of the tests in financial applications where samples are often many times larger. In such cases the power loss resulting from the estimation of a rich autocorrelation structure would also be of less importance. The sample size in our empirical application below is $n = 336$, so for the application we expect the performance of the tests to lie between the two cases considered in the present simulation study.

6 EXCHANGE RATE DYNAMICS

The analysis of exchange rate dynamics and potential (fractional) cointegrating relations between exchange rates for different currencies has attracted much attention recently. Baillie & Bollerslev (1989) find evidence of one cointegrating relation between seven different (log) spot exchange rates using conventional cointegration methods. This is challenged by Diebold, Gardeazabal & Yilmaz (1994) who show that the inclusion of an intercept changes the conclusion for the Baillie & Bollerslev (1989) data set. This finding is further supported in an analysis of a different data set covering a longer span of time in Diebold et al. (1994).

In the article by Baillie & Bollerslev (1994) it is argued that the failure of conventional cointegration tests to find evidence of cointegration in the Baillie & Bollerslev (1989) exchange rate data is due to the presence of fractional cointegration. Thus, they estimate the cointegration vector by OLS following Cheung & Lai (1993) and fit a simple fractionally integrated white noise model to the residuals. It is concluded that the exchange rates can be described by a $CI(1, .11)$ relationship (in our notation). However, their estimate of the integration order of the equilibrium errors (.89) may well be upwards biased since relevant short-run dynamics may have been left out. This is indeed what is concluded by Kim & Phillips (2001) who employ their fractional fully modified estimation procedure to a different data set covering a longer time span but the same exchange rates. They find that the equilibrium errors are best described by an ARFIMA(1,d,0) process with $d = .33$.

All the above studies concentrate on the estimation of the cointegration vector and the memory parameter of the equilibrium errors, but no formal testing of the hypothesis of fractional cointegration is attempted. We take the opposite view and concentrate on testing for the presence of (i) standard $I(1)-I(0)$ cointegration against fractional alternatives, (ii) $CI(\hat{d}, \hat{d})$ cointegration, where $\hat{d}$ is a preliminary estimate of d, and (iii) fractional cointegration with equilibrium errors that are integrated of order less than one-quarter. We apply our tests to a system of log exchange rates for the currencies of the following seven countries, (West) Germany, United Kingdom, Japan, Canada, France, Italy, and Switzerland against the US Dollar. The same currencies are examined in the studies cited above. However, where Baillie & Bollerslev (1989, 1994) and Diebold et al. (1994) consider daily observations covering 1 March 1980 to 28 January 1985 and Kim & Phillips (2001) consider quarterly observations from 1957 through 1997, our data set is comprised of monthly averages of noon (EST) buying rates and runs from January 1974 through December 2001 for a total of $n = 336$ observations. Thus, our data set, which is extracted from the Federal Reserve Board of Governors G.5 release, covers only the period of the current flexible exchange rate regime, but a much longer span of time than the Baillie & Bollerslev (1989) data set. A long time span has generally been found to be important in detecting long-run relations.

Table 5 about here

Table 5 presents the fractional integration analysis of the data set. The first two rows are the estimates of the fractional integration orders estimated by the conditional ML technique (CMLE) in Tanaka (1999) with lag orders $p = 0$ and $p = 1$. The standard errors reported in parenthesis are calculated as $\sqrt{6}/\pi$ when $p = 0$ and ω^{-1} when $p = 1$, where $\omega^2 = \pi^2/6 - \left(1 - a^2\right) a^{-1} \left(\ln\left(1 - a\right)\right)^2$ and a is the estimated AR coefficient, see Tanaka (1999). As a robustness check we also report the Gaussian semiparametric (GSP) estimates of Robinson (1995) with two different bandwidths in the final two rows. The standard errors of these estimates are $1/\left(2\sqrt{m}\right)$, see Robinson (1995). The final column gives estimates of a common integration order, computed simply as an average of the estimated integration orders for each exchange rate, which we use in our fractional cointegration analysis.

The estimates clearly show that the exchange rates can be well described as $I\left(1\right)$ processes. The CMLE estimates are insignificantly different from unity except CAN with $p = 0$, but that estimate may be upwards biased if relevant short-run dynamics is left out of the estimation. Thus, when $p = 1$ the CAN estimate is insignificantly different from unity. The GSP estimates are all insignificantly different from unity except FRA with $m = 67$. Hence, the results of Table 5 support the overwhelming evidence in the previous literature that exchange rates are $I(1)$. E.g. Baillie & Bollerslev (1989) conduct unit root tests of the $I(1)$ hypothesis against the $I(0)$ alternative and Baillie (1996) provides evidence from fractional models.

Table 6 about here

In Table 6 the results of the one-sided LM test (31) on the exchange rate data are presented. We consider all the different specifications outlined in Assumption 1 (with $p = 1$ in cases 1.1-1.3), and test three different hypotheses. Based on the evidence in Table 5 and the previous literature, we specify $d = b = 1$ in the first hypothesis corresponding to the standard $I(1) - I(0)$ model as discussed above. Secondly, we use the estimated common integration order $\hat{d}_c$ for d and b (i.e. we set $d = b = \hat{d}_c$). Under Assumptions 1.0 and 1.1 we use the estimates from Table 5 with $p = 0$ for $\hat{d}_c$, and under Assumptions 1.2 and 1.3 we use the estimates with $p = 1$. The third hypothesis, $d = 1, b = .76$, is that there exists a cointegrating relation which is integrated of order less than one-quarter (using $\varepsilon = .01$).

The results we obtain are mixed. Under Assumptions 1.0 and 1.2 all the tests reject strongly. However, when allowance is made for an autoregressive specification in the cointegrating relation, i.e. under Assumptions 1.1 and 1.3, the tests do not reject the third hypothesis (the third test under Assumption 1.3 has negative estimated Fisher information and is replaced by the approximation (33)), thus supporting a dynamic specification of the cointegrating relation. Indeed, under Assumption 1.3 none of the tests are significant. In the cases with an estimated autoregressive term in the equilibrium errors, the estimates of the autoregressive parameter (not reported in the table) are between .83 and .99, where .99 is the boundary we have chosen when implementing the tests to ensure stationarity of $g(z)$. Hence, there appears to be persistence in the cointegrating relation, but the results of Table 6 suggest that it could be only short memory.

7 CONCLUSION

We have proposed and examined a time domain LM test for the null of cointegration in a fractionally cointegrated model with the usual computational motivation. In the important case where the null hypothesis is that of standard $I(1) - I(0)$ cointegration, but the test is against fractional alternatives, the calculation of the LM test statistic does not require any fractional differencing and can be based on residuals from readily available computer software.

The likelihood theory in the time domain is tractable and the ML estimation of the cointegration vector β reduces to a version of the fully modified least squares estimator. Thus, the LM test statistic utilizes fully modified residuals to cancel the endogeneity and serial correlation biases. The test statistic is shown to have standard distributional properties under the null and under local alternatives, such that inference can be drawn from the normal and chi-squared distributions.

In the special case with *i.i.d.* Gaussian errors, the asymptotic Gaussian power envelope of all tests is achieved by the one-sided version of our test, and the asymptotic Gaussian power envelope of all unbiased tests is achieved by the two-sided version of our test. Thus, with *i.i.d.* Gaussian errors, the one-sided (two-sided) version of our test is uniformly most powerful among all (unbiased) tests.

The empirical relevance of our test is established by Monte Carlo experiments, which show that finite sample rejection frequencies are reasonable for samples of size $n = 200$ and close to the asymptotic local power for $n = 500$.

Finally, we have applied our methodology to the analysis of exchange rate dynamics in a system of exchange rates for seven major currencies against the US Dollar. We have focused on testing for the presence of (fractional) cointegration, rather than the estimation of any particular model, but the evidence is mixed.

ACKNOWLEDGEMENTS

This paper has benefitted from comments by Niels Haldrup, Michael Jansson, Peter Phillips, Katsumi Shimotsu, seminar participants at the University of Aarhus and Yale University, and two anonymous referees. It was completed while the author was visiting Yale University and the Cowles Foundation, their hospitality is gratefully acknowledged.

APPENDIX: PROOFS

Before we prove the theorems we need a lemma. Define the sample autocovariance and residual autocovariance functions

$$C(j) = \frac{1}{n} \sum_{t=j+1}^{n} e_t e'_{t-j} \text{ and } \hat{C}(j) = \frac{1}{n} \sum_{t=j+1}^{n} \hat{e}_t \hat{e}'_{t-j},$$

where the $\hat{e}_t$ are estimated residuals of a VAR(p) process. We consider the asymptotic distribution of a particular linear combination of the residual autocovariances in each of the four cases outlined in Assumption 1.

Lemma 1 *Let $\hat{e}_t$ be the estimated residuals of the K-dimensional VAR(p) process $A(L) u_t = e_t$, where e_t is i.i.d. $(0, \Sigma)$ with finite fourth moments and $A(z)$ has the structural parameterization in Assumption*

1.i. Then

$$\sqrt{n}\sum_{j=1}^{n-1} j^{-1}\operatorname{vec} C(j) \xrightarrow{D} N(0,\Omega_0)$$

$$\sqrt{n}\sum_{j=1}^{n-1} j^{-1}\operatorname{vec} \hat{C}(j) \xrightarrow{D} N(0,\Omega_i)$$

as $n \to \infty$, where

$$\Omega_0 = \frac{\pi^2}{6}\Sigma\otimes\Sigma,$$

$$\Omega_i = \frac{\pi^2}{6}\Sigma\otimes\Sigma - \left(\Sigma\Phi_{i1}'\otimes I_K,...,\Sigma\Phi_{ip}'\otimes I_K\right) H_i \left(H_i'\left(\Gamma_i\otimes\Sigma^{-1}\right)H_i\right)^{-1} H_i'\left(\Sigma\Phi_{i1}'\otimes I_K,...,\Sigma\Phi_{ip}'\otimes I_K\right)',$$

for $i = 1,2,3$. Here, Γ_i is the covariance matrix of $(u_t',...,u_{t-p+1}')'$, $\Phi_{il} = \sum_{j=l}^{\infty} j^{-1}\Psi_{i,j-l}$, $\Psi_{i,k}$ is the k'th term in the Wold representation of u_t normalized such that $\Psi_{i,0} = I_K$, and $H_i = \left(\partial a_1'/\partial\gamma_i,...,\partial a_p'/\partial\gamma_i\right)'$, where $a_j = \operatorname{vec} A_j$ are the coefficients in the autoregressive representation $A(L)u_t = e_t$.

Proof. For a fixed $m > p$ define the K^2m-vectors $C_m = \operatorname{vec}(C(1),...,C(m))'$ and $\hat{C}_m = \operatorname{vec}(\hat{C}(1),...,\hat{C}(m))'$. Consider first case 1.0, where u_t is *i.i.d.* and $C(j)$ is observable. It is well known that in this case

$$\sqrt{n}C_m \xrightarrow{D} N(0, I_m\otimes\Sigma\otimes\Sigma)$$

and thus

$$\sqrt{n}\sum_{j=1}^{m} j^{-1}\operatorname{vec} C(j) \xrightarrow{D} N\left(0,\sum_{j=1}^{m} j^{-2}\Sigma\otimes\Sigma\right).$$

The desired result for case 1.0 now follows by application of Bernstein's Lemma, see e.g. Hall & Heyde (1980, pp. 191-192).

For the remaining three cases we employ a result of Ahn (1988) on the asymptotic distribution of the residual autocovariances of a VAR(p) process under structural parameterization. Consider case 1.2. Define the matrix $H_2 = \left(\partial a_1'/\partial\gamma_2,...,\partial a_p'/\partial\gamma_2\right)'$, where $a_j = \operatorname{vec} A_j$ are the coefficients in the autoregressive representation $A(L)u_t = e_t$ and γ_2 is the vector of coefficients in $G(z)$. In this setup, Ahn (1988) showed that (in our notation)

$$\sqrt{n}\hat{C}_m \xrightarrow{D} N\left(0, I_m\otimes\Sigma\otimes\Sigma - G_m H_2\left(H_2'\left(\Gamma\otimes\Sigma^{-1}\right)H_2\right)^{-1} H_2' G_m'\right),$$

where

$$G_m' = \begin{bmatrix} \Sigma\otimes I_K & \Psi_1\Sigma\otimes I_K & \cdots & \cdots & \cdots & \Psi_{m-1}\Sigma\otimes I_K \\ 0 & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & & & \vdots \\ 0 & \cdots & 0 & \Sigma\otimes I_K & \cdots & \Psi_{m-p}\Sigma\otimes I_K \end{bmatrix}.$$

Consequently,

$$\sqrt{n}\sum_{j=1}^{m} j^{-1}\operatorname{vec}\hat{C}(j) \xrightarrow{D} N\left(0,\Omega_2^{(m)}\right),$$

where $\Omega_2^{(m)}$ is a truncated version of Ω_2, i.e. with $\pi^2/6$ replaced by $\sum_{j=1}^{m} j^{-2}$ and Φ_{il} replaced by $\Phi_{il}^{(m)}$, which is truncated at m. Again, we can apply Bernstein's Lemma to replace the truncated sums by their limits. For cases 1.1 and 1.3 the same results hold, except that H_i, Φ_{il}, and Γ_i are different as indicated by the subscript i. ■

As a simple example consider a bivariate VAR(1) system with $g(z) = 1 - \gamma_1 z$ and $G(z) = 1 - \gamma_2 z$. The H_i matrices are $H_1 = (1,0,0,0)'$, $H_2 = (0,0,0,1)'$, and $H_3 = (H_1, H_2)$ and the covariance equations simplify to

$$\Omega_i = \frac{\pi^2}{6}\Sigma\otimes\Sigma - \left(\Sigma\Phi_{i1}'\otimes I_K\right)H_i\left(H_i'\left(\Sigma^{-1}\otimes\Gamma_i\right)H_i\right)^{-1}H_i'\left(\Sigma\Phi_{i1}\otimes I_K\right),\ i=1,2,3,$$

where $\Phi_{11} = \operatorname{diag}(\phi_1, 1)$, $\Phi_{21} = \operatorname{diag}(1, \phi_2)$, $\Phi_{31} = \operatorname{diag}(\phi_1, \phi_2)$, $\phi_i = -\gamma_i^{-1}\ln(1-\gamma_i)$, $i = 1, 2$, and $\Gamma_i = E(u_t u_t')$ can be estimated by $n^{-1}\sum_{t=1}^{n}\hat{u}_t\hat{u}_t'$.

Proof of Theorem 3.1. Suppose first that β is known. Using that $\operatorname{vec}(A)'\operatorname{vec}(B) = \operatorname{tr}(A'B)$ and by application of Lemma 1, the score statistic is

$$\begin{aligned} S_n &= \sqrt{n}\sum_{j=1}^{n-1} j^{-1}\operatorname{vec}\left(\Sigma^{-1}e_1e_1'\right)'\operatorname{vec}\left(\hat{C}(j)\right) \\ &\xrightarrow{D} N\left(0, \operatorname{vec}\left(\Sigma^{-1}e_1e_1'\right)'\Omega_i\operatorname{vec}\left(\Sigma^{-1}e_1e_1'\right)\right), \end{aligned}$$

where the Ω_i are defined in Lemma 1. The variance equations (26) and (27) follow immediately from Lemma 1, e.g. in case 1.0 with *i.i.d.* errors the variance is

$$\begin{aligned} \operatorname{vec}\left(\Sigma^{-1}e_1e_1'\right)'\frac{\pi^2}{6}(\Sigma\otimes\Sigma)\operatorname{vec}\left(\Sigma^{-1}e_1e_1'\right) &= \frac{\pi^2}{6}\operatorname{vec}\left(\Sigma^{-1}e_1e_1'\right)'\operatorname{vec}\left(e_1e_1'\Sigma\right) \\ &= \frac{\pi^2}{6}\operatorname{tr}\left(e_1e_1'\Sigma^{-1}e_1e_1'\Sigma\right) \\ &= \frac{\pi^2}{6}\frac{\sigma_{11}^2}{\sigma_{1.2}^2}. \end{aligned}$$

Next, we show that estimating β does not influence the result. From e.g. Cheung & Lai (1993), Marinucci & Robinson (2001), and Kim & Phillips (2001) we know that, since $\hat{\beta}$ is estimated by OLS between $I(b)$ processes with $I(0)$ errors, $\hat{\beta} - \beta = O_p(n^{1-2b})$ when $b \le 1$ and $\hat{\beta} - \beta = O_p(n^{-b})$ when $b > 1$.

For simplicity we consider only the case with *i.i.d.* errors in the remainder of the proof, i.e. $u_t = e_t$, the general case follows similarly. Consider the residual processes

$$\begin{aligned}
\hat{z}_t &= y_{1t} - \hat{\beta}' y_{2t} = z_t + (\beta - \hat{\beta})' y_{2t}, \\
\hat{u}_{1t} &= \Delta^{d-b} \hat{z}_t = u_{1t} + \Delta^{d-b} (\beta - \hat{\beta})' y_{2t}, \\
\hat{u}_{1.2t} &= u_{1.2t} + \Delta^{d-b} (\beta - \hat{\beta})' y_{2t},
\end{aligned}$$

and define $w_t = \sum_{j=1}^{m} j^{-1} u_{1t-j} u_{1.2t}$ and $\hat{w}_t = \sum_{j=1}^{m} j^{-1} \hat{u}_{1t-j} \hat{u}_{1.2t}$. Then

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\hat{w}_t - w_t) &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sum_{j=1}^{m} j^{-1} \Big((u_{1t-j} + \Delta^{d-b}(\beta - \hat{\beta})' y_{2t-j})(u_{1.2t} + \Delta^{d-b}(\beta - \hat{\beta})' y_{2t}) - u_{1t-j} u_{1.2t} \Big) \\
&= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sum_{j=1}^{m} j^{-1} \Big((\Delta^{d-b}(\beta - \hat{\beta})' y_{2t-j})(\Delta^{d-b}(\beta - \hat{\beta})' y_{2t}) \\
&\quad + u_{1t-j} \Delta^{d-b} (\beta - \hat{\beta})' y_{2t} + u_{1.2t} \Delta^{d-b} (\beta - \hat{\beta})' y_{2t-j} \Big) \\
&= O_p \left(\sum_{j=1}^{m} j^{-1} \left[\frac{1}{\sqrt{n}} \sum_{t=j+1}^{n} \Big((\beta - \hat{\beta})' \Delta^{-b} u_{2t-j} \Big) \Big((\beta - \hat{\beta})' \Delta^{-b} u_{2t} \Big) \right. \right. \\
&\quad \left. \left. + \frac{1}{\sqrt{n}} \sum_{t=j+1}^{n} u_{1t-j} (\beta - \hat{\beta})' \Delta^{-b} u_{2t} + \frac{1}{\sqrt{n}} \sum_{t=j+1}^{n} u_{1.2t} (\beta - \hat{\beta})' \Delta^{-b} u_{2t-j} \right] \right).
\end{aligned}$$

When $b \leq 1$ the first term is

$$O_p \left(n^{3/2-4b} \sum_{t=j+1}^{n} (\Delta^{-b} u_{2t-j})(\Delta^{-b} u_{2t})' \right) = O_p(n^{3/2-2b}),$$

and when $b > 1$ it is $O_p\left(n^{-1/2}\right)$. Similarly, when $b \leq 1$ the second and third terms are of orders $O_p\left(n^{1/2-2b} \sum_{t=j+1}^{n} u_{1t-j} \Delta^{-b} u_{2t}\right) = O_p(n^{1/2-b})$ and $O_p\left(n^{1/2-2b} \sum_{t=j+1}^{n} u_{1.2t} \Delta^{-b} u_{2t}\right) = O_p(n^{1/2-b})$, respectively, and when $b > 1$ they are both $O_p\left(n^{-1/2}\right)$. Since $b > 3/4$ by assumption, all these terms are $o_p(1)$ and we are done. ■

Proof of Theorem 3.2. As in the proof of Theorem 3.1 it can be shown that estimating β does not affect the result, so we assume that β is known. The second derivative of the likelihood (18) is

$$\begin{aligned}
\frac{\partial^2 L(\theta, \beta, \Sigma)}{\partial \theta^2} &= -\frac{1}{\sigma_{1.2}^2} \sum_{t=1}^{n} \left\{ \ln(1-L) \ln(1-L) \left(g(L) \Delta^{d-b+\theta} \left(y_{1t} - \beta' y_{2t} \right) \right) \right\} \\
&\quad \times \left(\left(g(L) \Delta^{d-b+\theta} \left(y_{1t} - \beta' y_{2t} \right) - \sigma_{21}' \Sigma_{22}^{-1} G(L) \Delta^{d} y_{2t} \right) \right) \\
&\quad - \frac{1}{\sigma_{1.2}^2} \sum_{t=1}^{n} \left\{ \ln(1-L) \left(g(L) \Delta^{d-b+\theta} \left(y_{1t} - \beta' y_{2t} \right) \right) \right\}^2 \\
&= -\frac{1}{\sigma_{1.2}^2} \sum_{t=1}^{n} \sum_{j=1}^{t-1} \sum_{k=1}^{t-j-1} j^{-1} k^{-1} \hat{e}_{1t-j-k} \hat{e}_{1.2t} - \frac{1}{\sigma_{1.2}^2} \sum_{t=1}^{n} \sum_{j=1}^{t-1} \sum_{k=1}^{t-1} j^{-1} k^{-1} \hat{e}_{1t-j} \hat{e}_{1t-k}. \quad (39)
\end{aligned}$$

In the case with *i.i.d.* errors, e_t is observable and the contribution of the first term to the Fisher information is zero by uncorrelatedness of e_t. The contribution of the second term is

$$E\frac{1}{n\sigma_{1.2}^2}\sum_{t=1}^{n}\sum_{j=1}^{t-1}\sum_{k=1}^{t-1}j^{-1}k^{-1}e_{1t-j}e_{1t-k}=\frac{1}{\sigma_{1.2}^2}\sum_{j=1}^{n-1}j^{-2}\sigma_{11}^2$$

by uncorrelatedness of e_t, which proves the result for *i.i.d.* errors.

In the remaining cases, we need to take the estimation of the autoregressive parameters into account. Again it can be shown that the first term of (39) is negligible. Since $\hat{C}(0)=\Sigma+O_p(n^{-1/2})$ and $\sigma_{1.2}^{-2}=e_1'\Sigma^{-1}e_1$, the contribution of the second term is

$$\begin{aligned}&E\,\mathrm{tr}\,\frac{1}{n}\sum_{t=1}^{n}\sum_{j=1}^{t-1}\sum_{k=1}^{t-1}j^{-1}k^{-1}e_1e_1'\Sigma^{-1}\hat{C}(0)\Sigma^{-1}\hat{e}_{1t-j}\hat{e}_{1t-k}\\ =\;&E\,\mathrm{tr}\,n\sum_{j=1}^{n-1}\sum_{k=1}^{n-1}j^{-1}k^{-1}e_1e_1'\Sigma^{-1}\hat{C}(j)\,e_1e_1'\hat{C}(k)'\,\Sigma^{-1}\\ =\;&En\sum_{j=1}^{n-1}j^{-1}e_1'\Sigma^{-1}\hat{C}(j)\,e_1\sum_{k=1}^{n-1}k^{-1}e_1'\Sigma^{-1}\hat{C}(k)\,e_1,\end{aligned}$$

which is equal to $\mathrm{vec}\left(\Sigma^{-1}e_1e_1'\right)'\Omega_i\,\mathrm{vec}\left(\Sigma^{-1}e_1e_1'\right)$ as $n\to\infty$ by Lemma 1 and Theorem 3.1. ■

Proof of Theorem 3.3. Let $\theta=\delta/\sqrt{n}$. First suppose β is known and define $\hat{e}_{1nt}=\hat{g}(L)\,\Delta^{d-b}z_t=\hat{g}(L)\,\Delta^{-\theta}e_{1t}$ and $\hat{e}_{2t}=\hat{G}(L)\,\Delta^{d}y_{2t}$. By the Mean Value Theorem we obtain

$$\hat{e}_{1nt}=\hat{e}_{1t}+\frac{\delta}{\sqrt{n}}\sum_{j=1}^{t-1}j^{-1}\hat{e}_{1t-j}+o_p(n^{-1/2})$$

for all $t=1,...,n$. Thus, under $\theta=\delta/\sqrt{n}$,

$$\begin{aligned}S_n\;&=\;\frac{1}{\sigma_{1.2}^2\sqrt{n}}\sum_{t=1}^{n}\sum_{j=1}^{t-1}j^{-1}\hat{e}_{1nt}(\hat{e}_{1nt}-\sigma_{21}'\Sigma_{22}^{-1}\hat{e}_{2t})\\ &=\;\frac{1}{\sigma_{1.2}^2\sqrt{n}}\sum_{t=1}^{n}\sum_{j=1}^{t-1}j^{-1}\left(\hat{e}_{1t-j}+\frac{\delta}{\sqrt{n}}\sum_{k=1}^{t-j-1}k^{-1}\hat{e}_{1t-j-k}\right)\left(\hat{e}_{1.2t}+\frac{\delta}{\sqrt{n}}\sum_{k=1}^{t-1}k^{-1}\hat{e}_{1t-k}\right)+o_p(1)\\ &=\;\sqrt{n}\sum_{j=1}^{n-1}j^{-1}e_1'\Sigma^{-1}\hat{C}(j)\,e_1+\delta n\sum_{j=1}^{n-1}j^{-1}e_1'\Sigma^{-1}\hat{C}(j)\,e_1\sum_{k=1}^{n-1}k^{-1}e_1'\Sigma^{-1}\hat{C}(k)\,e_1+o_p(1)\end{aligned}$$

as in the proofs of Theorems 3.1 and 3.2. The result when β is known now follows from Lemma 1 and the above theorems. When β is unknown we can apply the same arguments as in the proof of Theorem 3.1, along with elementary inequalities to the components due to $\hat{e}_{1nt}-\hat{e}_{1t}$, to show that the result is unaffected. ■

Proof of Corollary 3.1. Follows immediately from Theorem 3.3. ■

Proof of Corollary 3.2. Follows immediately from (32) and Theorem 3.2. ■

Proof of Theorem 4.1. By the Mean Value Theorem we obtain

$$\begin{aligned}\tilde{u}_{1.2nt} &= u_{1t} - \sigma_{21}'\Sigma_{22}^{-1}u_{2t} + \frac{c}{\sqrt{n}}\sum_{j=1}^{t-1} j^{-1}u_{1t-j} + o_p(n^{-1/2}) \\ \hat{u}_{1.2nt} &= u_{1t} - \sigma_{21}'\Sigma_{22}^{-1}u_{2t} + \frac{c-\delta}{\sqrt{n}}\sum_{j=1}^{t-1} j^{-1}u_{1t-j} + o_p(n^{-1/2})\end{aligned}$$

for all $t = 1, ..., n$. Thus, we note that

$$\frac{1}{n}\sum_{t=1}^{n}\tilde{u}_{1.2nt}^2 \xrightarrow{P} \sigma_{1.2}^2 \tag{40}$$

as $n \to \infty$. The numerator of M_n is

$$\begin{aligned}\sum_{t=1}^{n}\tilde{u}_{1.2nt}^2 - \sum_{t=1}^{n}\hat{u}_{1.2nt}^2 &= \sum_{t=1}^{n}\frac{c^2}{n}\sum_{j=1}^{t-1} j^{-2}u_{1t-j}^2 - \sum_{t=1}^{n}\frac{(c-\delta)^2}{n}\sum_{j=1}^{t-1} j^{-2}u_{1t-j}^2 \\ &\quad +2\sum_{t=1}^{n}u_{1.2t}\frac{\delta}{\sqrt{n}}\sum_{j=1}^{t-1} j^{-1}u_{1t-j} + o_p(1) \\ &= \delta(2c-\delta)\frac{\pi^2}{6}\sigma_{11}^2 + 2\delta\sqrt{\frac{\pi^2}{6}\sigma_{11}^2\sigma_{1.2}^2}Z + o_p(1). \end{aligned} \tag{41}$$

Combining (40) and (41) we get the desired result. ■

Proof of Theorem 4.2. Consider the one-sided case with $\delta > 0$ (the reverse case follows similarly). The one-sided power envelope is

$$\begin{aligned}\Pi(\delta) &= P(M(\delta,\delta) > c_\alpha(\delta)) \\ &= P\left(\delta\left(2\sqrt{\mathcal{I}_0}Z + \delta\mathcal{I}_0\right) > c_\alpha(\delta)\right) \\ &= P\left(Z > \left(\frac{c_\alpha(\delta)}{\delta} - \delta\mathcal{I}_0\right)/2\sqrt{\mathcal{I}_0}\right),\end{aligned}$$

where $c_\alpha(\delta)$ satisfies

$$\begin{aligned}\alpha &= P(M(0,\delta) > c_\alpha(\delta)) \\ &= P\left(Z > \left(\frac{c_\alpha(\delta)}{\delta} + \delta\mathcal{I}_0\right)/2\sqrt{\mathcal{I}_0}\right)\end{aligned}$$

such that $c_\alpha(\delta) = 2\delta\sqrt{\mathcal{I}_0}Z_\alpha - \delta^2\mathcal{I}_0$. Then

$$\begin{aligned}\Pi(\delta) &= P\left(Z > \left(2\sqrt{\mathcal{I}_0}Z_\alpha - 2\delta\mathcal{I}_0\right)/2\sqrt{\mathcal{I}_0}\right) \\ &= P\left(Z > Z_\alpha - \delta\sqrt{\mathcal{I}_0}\right).\end{aligned}$$

In the two-sided case we note that, since for varying c the family of distributions $M(c,\delta)$ is normal, it satisfies the requirement that it be strictly totally positive of order 3 (STP_3, see Lehmann (1986, p. 119)). Hence the power envelope of all unbiased tests of $H_0:\theta=0$ against $H_1:\theta_{1n}=\delta/\sqrt{n}$ is given by $\Pi_2(\delta)=1-P(C_{1,\alpha}(\delta)<M(\delta,\delta)<C_{2,\alpha}(\delta))$ (Lehmann (1986, p. 303)), where the constants are determined by

$$P(C_{1,\alpha}(\delta)<M(0,\delta)<C_{2,\alpha}(\delta)) = 1-\alpha \quad (42)$$

$$\left.\frac{\partial P(C_{1,\alpha}(\delta)<M(c,\delta)<C_{2,\alpha}(\delta))}{\partial c}\right|_{c=0} = 0. \quad (43)$$

Consider first (43) which implies that ($\phi(\cdot)$ is the density function of the standard normal distribution)

$$\phi\left(\frac{C_{2,\alpha}(\delta)+\delta^2\mathcal{I}_0}{2\delta\sqrt{\mathcal{I}_0}}\right)=\phi\left(\frac{C_{1,\alpha}(\delta)+\delta^2\mathcal{I}_0}{2\delta\sqrt{\mathcal{I}_0}}\right)$$

with the non-trivial solution $C_{1,\alpha}(\delta)=-C_{2,\alpha}(\delta)-2\delta^2\mathcal{I}_0$. Now we can find the constants from (42),

$$\begin{aligned}1-\alpha &= P\left(-C_{2,\alpha}(\delta)-2\delta^2\mathcal{I}_0<M(0,\delta)<C_{2,\alpha}(\delta)\right)\\ &= P\left(-\frac{C_{2,\alpha}(\delta)+\delta^2\mathcal{I}_0}{2\delta\sqrt{\mathcal{I}_0}}<Z<\frac{C_{2,\alpha}(\delta)+\delta^2\mathcal{I}_0}{2\delta\sqrt{\mathcal{I}_0}}\right),\end{aligned}$$

where Z is a standard normal random variable. Thus, $C_{2,\alpha}(\delta)$ is the solution to $\Phi\left(\left(C_{2,\alpha}(\delta)+\delta^2\mathcal{I}_0\right)/2\delta\sqrt{\mathcal{I}_0}\right)=1-\alpha/2$, which implies $C_{2,\alpha}(\delta)=2\delta\sqrt{\mathcal{I}_0}Z_{1-\alpha/2}-\delta^2\mathcal{I}_0$, where $Z_{1-\alpha/2}$ is the $100(1-\alpha/2)\%$ point of the standard normal distribution.

The two-sided power envelope is then

$$\begin{aligned}\Pi_2(\delta) &= 1-P(C_{1,\alpha}(\delta)<M(\delta,\delta)<C_{2,\alpha}(\delta))\\ &= 1-P\left(-2\delta\sqrt{\mathcal{I}_0}Z_{1-\alpha/2}-\delta^2\mathcal{I}_0<2\delta\sqrt{\mathcal{I}_0}Z+\delta^2\mathcal{I}_0<2\delta\sqrt{\mathcal{I}_0}Z_{1-\alpha/2}-\delta^2\mathcal{I}_0\right)\\ &= 1-P\left(-Z_{1-\alpha/2}<Z+\delta\sqrt{\mathcal{I}_0}<Z_{1-\alpha/2}\right)\\ &= 1-F_\lambda\left(\chi^2_{1,1-\alpha}\right).\end{aligned}$$

■

References

Agiakloglou, C. & Newbold, P. (1994), 'Lagrange multiplier tests for fractional difference', *Journal of Time Series Analysis* **15**, 253–262.

Ahn, S. K. (1988), 'Distribution for residual autocovariances in multivariate autoregressive models with structural parameterization', *Biometrika* **75**, 590–593.

Baillie, R. T. (1996), 'Long memory processes and fractional integration in econometrics', *Journal of Econometrics* **73**, 5–59.

Baillie, R. T. & Bollerslev, T. (1989), 'Common stochastic trends in a system of exchange rates', *Journal of Finance* **44**, 167–181.

Baillie, R. T. & Bollerslev, T. (1994), 'Cointegration, fractional cointegration, and exchange rate dynamics', *Journal of Finance* **49**, 737–745.

Cheung, Y. E. & Lai, K. S. (1993), 'A fractional cointegration analysis of purchasing power parity', *Journal of Business and Economic Statistics* **11**, 103–122.

Diebold, F. X., Gardeazabal, J. & Yilmaz, K. (1994), 'On cointegration and exchange rate dynamics', *Journal of Finance* **49**, 727–735.

Diebold, F. X. & Rudebusch, G. D. (1989), 'Long memory and persistence in aggregate output', *Journal of Monetary Economics* **24**, 189–209.

Doornik, J. A. (2001), *Ox: An Object-Oriented Matrix Language*, 4th edn, Timberlake Consultants Press, London.

Doornik, J. A. & Ooms, M. (2001), 'A package for estimating, forecasting and simulating arfima models: Arfima package 1.01 for Ox', *Working Paper, Nuffield College, Oxford* .

Dueker, M. & Startz, R. (1998), 'Maximum-likelihood estimation of fractional cointegration with an application to U.S. and Canadian bond rates', *Review of Economics and Statistics* **83**, 420–426.

Elliott, G., Rothenberg, T. J. & Stock, J. H. (1996), 'Efficient tests for an autoregressive unit root', *Econometrica* **64**, 813–836.

Engle, R. & Granger, C. W. J. (1987), 'Cointegration and error correction: Representation, estimation and testing', *Econometrica* **55**, 251–276.

Fox, R. & Taqqu, M. S. (1986), 'Large-sample properties of parameter estimates for strongly dependent stationary gaussian series', *Annals of Statistics* **14**, 517–532.

Granger, C. W. J. (1981), 'Some properties of time series data and their use in econometric model specification', *Journal of Econometrics* **16**, 121–130.

Hall, P. & Heyde, C. C. (1980), *Martingale Limit Theory and its Application*, Academic Press, New York.

Hosking, J. R. M. (1980), 'The multivariate portmanteau statistic', *Journal of the American Statistical Association* **75**, 602–608.

Jansson, M. (2001), 'Point optimal tests of the null hypothesis of cointegration', *Preprint, University of California, Berkeley* .

Jansson, M. & Haldrup, N. (2001), 'Regression theory for nearly cointegrated time series', *Forthcoming in Econometric Theory* .

Jeganathan, P. (1999), 'On asymptotic inference in cointegrated time series with fractionally integrated errors', *Econometric Theory* **15**, 583–621.

Kim, C. S. & Phillips, P. C. B. (2001), 'Fully modified estimation of fractional cointegration models', *Preprint, Yale University* .

Lehmann, E. L. (1986), *Testing Statistical Hypotheses*, 2nd edn, Springer, New York.

Lobato, I. N. & Velasco, C. (2000), 'Long memory in stock-market trading volume', *Journal of Business and Economic Statistics* **18**, 410–427.

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. (2001), 'Semiparametric fractional cointegration analysis', *Journal of Econometrics* **105**, 225–247.

Nielsen, M. Ø. (2001), 'Efficient likelihood inference in nonstationary univariate models', *Department of Economics Working Paper 2001-08, University of Aarhus* .

Phillips, P. C. B. (1991), 'Optimal inference in cointegrated systems', *Econometrica* **59**, 283–306.

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

Phillips, P. C. B. & Loretan, M. (1991), 'Estimating long-run economic equilibria', *Review of Economic Studies* **58**, 407–436.

Robinson, P. M. (1991), 'Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regressions', *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.

Saikkonen, P. (1991), 'Asymptotically efficient estimation of cointegration regressions', *Econometric Theory* **7**, 1–21.

Shin, Y. (1994), 'A residual-based test of the null of cointegration against the alternative of no cointegration', *Econometric Theory* **10**, 91–115.

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.

Table 1: Finite Sample Rejection Frequencies Under Assumption 1.0

		Uncorrelated			Correlation .6		
Sample Size	θ	Envelope	LM	LMsc	Envelope	LM	LMsc
$n = 200$	0	0.050	0.024	0.050	0.050	0.031	0.050
	0.05	0.230	0.172	0.247	0.305	0.222	0.275
	0.10	0.567	0.468	0.573	0.733	0.585	0.643
	0.15	0.859	0.755	0.828	0.960	0.867	0.886
	0.20	0.976	0.913	0.949	0.998	0.976	0.985
	0.25	0.998	0.986	0.993	1.000	0.996	0.998
$n = 500$	0	0.050	0.037	0.050	0.050	0.040	0.050
	0.05	0.416	0.360	0.416	0.559	0.507	0.555
	0.10	0.889	0.838	0.879	0.974	0.944	0.957
	0.15	0.996	0.980	0.986	1.000	0.997	0.997
	0.20	1.000	1.000	1.000	1.000	1.000	1.000
	0.25	1.000	1.000	1.000	1.000	1.000	1.000

Table 2: Finite Sample Rejection Frequencies Under Assumption 1.1

		Uncorrelated			Correlation .6		
Sample Size	θ	Envelope	LM	LMsc	Envelope	LM	LMsc
$n = 200$	0	0.050	0.022	0.050	0.050	0.024	0.050
	0.05	0.121	0.060	0.117	0.146	0.075	0.109
	0.10	0.243	0.116	0.195	0.323	0.136	0.197
	0.15	0.412	0.191	0.289	0.553	0.229	0.305
	0.20	0.600	0.270	0.383	0.766	0.330	0.415
	0.25	0.766	0.348	0.455	0.906	0.397	0.469
$n = 500$	0	0.050	0.029	0.050	0.050	0.031	0.050
	0.05	0.185	0.123	0.159	0.240	0.161	0.208
	0.10	0.442	0.261	0.336	0.591	0.356	0.432
	0.15	0.727	0.464	0.543	0.878	0.594	0.664
	0.20	0.912	0.611	0.673	0.982	0.762	0.810
	0.25	0.982	0.742	0.802	0.999	0.851	0.887

Table 3: Finite Sample Rejection Frequencies Under Assumption 1.2

		Uncorrelated			Correlation .6		
Sample Size	θ	Envelope	LM	LMsc	Envelope	LM	LMsc
$n = 200$	0	0.050	0.017	0.050	0.050	0.015	0.050
	0.05	0.230	0.178	0.312	0.286	0.146	0.226
	0.10	0.567	0.452	0.581	0.696	0.421	0.580
	0.15	0.859	0.742	0.836	0.944	0.666	0.798
	0.20	0.976	0.922	0.956	0.996	0.903	0.952
	0.25	0.998	0.975	0.992	1.000	0.972	0.991
$n = 500$	0	0.050	0.035	0.050	0.050	0.023	0.050
	0.05	0.416	0.378	0.436	0.524	0.308	0.444
	0.10	0.889	0.819	0.850	0.961	0.809	0.880
	0.15	0.996	0.987	0.993	1.000	0.983	0.995
	0.20	1.000	1.000	1.000	1.000	1.000	1.000
	0.25	1.000	1.000	1.000	1.000	1.000	1.000

Table 4: Finite Sample Rejection Frequencies Under Assumption 1.3

		Uncorrelated			Correlation .6		
Sample Size	θ	Envelope	LM	LMsc	Envelope	LM	LMsc
$n = 200$	0	0.050	0.022	0.050	0.050	0.024	0.050
	0.05	0.121	0.058	0.088	0.146	0.065	0.110
	0.10	0.243	0.115	0.193	0.321	0.148	0.233
	0.15	0.412	0.223	0.319	0.550	0.251	0.356
	0.20	0.600	0.269	0.358	0.763	0.321	0.419
	0.25	0.766	0.332	0.432	0.904	0.355	0.478
$n = 500$	0	0.050	0.029	0.050	0.050	0.027	0.050
	0.05	0.185	0.122	0.177	0.238	0.148	0.239
	0.10	0.442	0.287	0.385	0.588	0.368	0.496
	0.15	0.727	0.483	0.598	0.876	0.599	0.722
	0.20	0.912	0.621	0.727	0.982	0.764	0.847
	0.25	0.982	0.734	0.805	0.999	0.836	0.889

Table 5: Estimates of Fractional Integration Orders

		WG (y_{1t})	CAN	SW	FRA	ITA	JAP	UK	$\hat{d}_c = \overline{\hat{d}}$
CMLE	$p = 0$	1.0057	1.1211**	0.9938	1.0081	1.0033	0.9975	1.0770	1.0295
		(0.0425)	(0.0425)	(0.0425)	(0.0425)	(0.0425)	(0.0425)	(0.0425)	
	$p = 1$	0.9625	1.0588	0.9465	0.9920	1.0023	0.9959	1.0163	0.9963
		(0.0975)	(0.0744)	(0.1016)	(0.0914)	(0.1026)	(0.0980)	(0.0960)	
GSP	$m = 33$	1.0428	1.1311	1.0197	1.1141	1.0857	0.9696	0.9837	1.0495
		(0.0870)	(0.0870)	(0.0870)	(0.0870)	(0.0870)	(0.0870)	(0.0870)	
	$m = 67$	1.0847	1.0406	1.0662	1.1338*	1.1029	1.1185	1.0725	1.0885
		(0.0611)	(0.0611)	(0.0611)	(0.0611)	(0.0611)	(0.0611)	(0.0611)	

Standard errors are given in parenthesis. One asterisk denotes significantly different from unity at 5% level and two asterisks denote significantly different from unity at 1% level.

Table 6: One-sided LM Tests for Fractional Cointegration

	LM_{01}	LM_{11}	LM_{21}	LM_{31}
$d = b = 1$	27.72**	3.509**	27.43**	0.7701
$d = b = \hat{d}_c$	27.82**	3.951**	27.31**	0.6280
$d = 1, b = 0.76$	24.92**	−1.487	23.72**	0.0242†

One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level. A dagger means the LM statistic did not compute and was replaced by the approximation $\widehat{LM}_{i1}$ in (33).

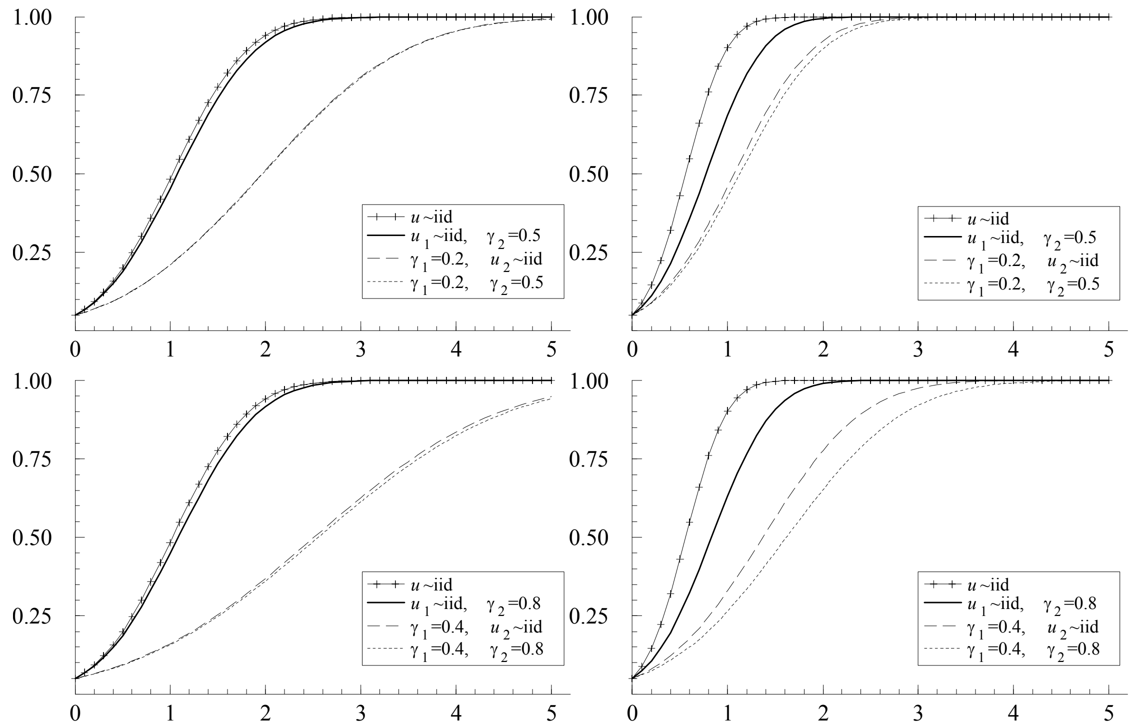

Figure 1: Asymptotic local power functions calculated using Corollary 3.1 with $d = b = 1$ and first order autoregressive specifications. The variances are normalized to unity, and the correlation is .6 and .9 in the left-hand and right-hand side panels, respectively.

Working Paper

2001-6:	Boriss Siliverstovs: Multicointegration in US consumption data.
2001-7:	Jakob Roland Munch and Michael Svarer: Rent Control and Tenancy Duration.
2001-8:	Morten Ø. Nielsen: Efficient Likelihood Inference in Non-stationary Univariate Models.
2001-9:	Effrosyni Diamantoudi: Stable Cartels Revisited.
2001-16:	Bjarne Brendstrup, Svend Hylleberg, Morten Nielsen, Lars Skipper and Lars Stentoft: Seasonality in Economic Models.
2001-17:	Martin Paldam: The Economic Freedom of Asian Tigers - an essay on controversy.
2001-18:	Celso Brunetti and Peter Lildholt: Range-based covariance estimation with a view to foreign exchange rates.
2002-1:	Peter Jensen, Michael Rosholm and Mette Verner: A Comparison of Different Estimators for Panel Data Sample Selection Models.
2002-2:	Torben M. Andersen: International Integration and the Welfare State.
2002-3:	Bo Sandemann Rasmussen: Credibility, Cost of Reneging and the Choice of Fixed Exchange Rate Regime.
2002-4:	Bo William Hansen and Lars Mayland Nielsen: Can Nominal Wage and Price Rigidities Be Equivalent Propagation Mechanisms? The Case of Open Economies.
2002-5:	Anna Christina D'Addio and Michael Rosholm: Left-Censoring in Duration Data: Theory and Applications.
2002-6:	Morten Ørregaard Nielsen: Efficient Inference in Multivariate Fractionally Integration Models.
2002-7:	Morten Ørregaard Nielsen. Optimal Residual Based Tests for Fractional Cointegration and Exchange Rate Dynamics.