Facultad de Ciencias Económicas y Empresariales
Universidad de Navarra

Working Paper nº 08/02

Root-n-Consistent Estimation of Weak Fractional Cointegration

Javier Hualde
Peter M. Robinson

Root-n-Consistent Estimation of Weak Fractional Cointegration
J. Hualde and P.M. Robinson
Working Paper No.08/02
November 2002
JEL Codes: C22

ABSTRACT

Empirical evidence has emerged of the possibility of fractional cointegration such that the gap, β, between the integration order δ of the observables and the integration order γ of the cointegrating errors is less than 0.5. This includes circumstances both when the observables are stationary or asymptotically stationary with long memory (so δ <1/2) and when they are nonstationary (so $\delta \geq$ 1/2). We call this weak cointegration, and it contrasts strongly with the traditional econometric prescription of unit root observables and short memory cointegrating errors, where β =1. Asymptotic inferential theory also differs from this case, and from other members of the class β >1/2, in particular $\sqrt{n}$-consistent and asymptotically normal estimation of the cointegrating vector ν is possible when β <1/2, as we explore in a simple bivariate model. The estimate depends on γ and δ or, more realistically, on estimates of unknown γ and δ. These latter estimates need to be $\sqrt{n}$-consistent, and the asymptotic distribution of the estimate of ν is sensitive to their precise form. We propose estimates of γ and δ that are computationally relatively convenient, relying on only univariate nonlinear optimization. Finite sample performance of the methods is examined by means of Monte Carlo simulations, and several applications to empirical data included.

Keywords: Fractional Cointegration; Parametric Estimation; Asymptotic Normality.

Javier Hualde
Universidad de Navarra
Facultad de Económicas
Campus Universitario
31080 Pamplona, Spain
jhualde@unav.es

Peter M. Robinson
London School of Economics
Department of Economics
Houghton Street
London WC2A 2AE, UK
p.m.robinson@lse.ac.uk

ACKNOWLEDGMENTS

Research supported by ESRC Grant R000238212. The first author's research was also supported by the Fundación Ramón Areces (Spain) and the second author's research was also supported by a Leverhulme Trust Personal Research Professorship.

1. Introduction

Cointegration analysis has usually proceeded under the assumption of unit root ($I(1)$) observable series and short-memory stationary ($I(0)$) cointegrating errors. Here, the least squares estimate (LSE) of the cointegrating vector is not only consistent, but super-consistent (with convergence rate equal to sample size, n) (Stock, 1987) despite contemporaneous correlation between regressors and cointegrating errors; optimal estimates, which account for this correlation, enjoy no better rate of convergence (Phillips, 1991).

It is also possible to consider cointegration in a fractional context. To be specific, we consider the model

$$\left.\begin{aligned} \Delta^{\gamma}(y_t - \nu x_t) = u_{1t}^{\#}, \quad t \geq 1, \quad y_t = 0, \quad t \leq 0, \\ \Delta^{\delta} x_t = u_{2t}^{\#}, \quad t \geq 1, \quad x_t = 0, \quad t \leq 0, \end{aligned}\right\} \tag{1}$$

for the bivariate observable sequence $\{y_t, x_t\}$. Here $\Delta = 1 - L$, where L is the lag operator;

$$(1-L)^{-\alpha} = \sum_{j=0}^{\infty} a_j(\alpha) L^j, \quad a_j(\alpha) = \frac{\Gamma(j+\alpha)}{\Gamma(\alpha)\Gamma(j+1)}, \tag{2}$$

taking $\Gamma(\alpha) = \infty$ for $\alpha = 0, -1, -2, ...$, and $\Gamma(0)/\Gamma(0) = 1$; the # superscript attached to a scalar or vector sequence v_t has the meaning

$$v_t^{\#} = v_t 1(t > 0), \tag{3}$$

where $1(\cdot)$ is the indicator function; $\{(u_{1t}, u_{2t}), t = 0, \pm 1, ...\}$ is an unobservable covariance stationary bivariate sequence having spectral density matrix that is nonsingular and bounded at all frequencies; and the real numbers γ and δ satisfy

$$0 \leq \gamma < \delta. \tag{4}$$

On this basis, we refer to $u_t = (u_{1t}, u_{2t})'$ as $I(0)$, x_t as $I(\delta)$ and $y_t - \nu x_t$ as $I(\gamma)$, while for

$$\nu \neq 0, \tag{5}$$

(4) implies that y_t is also $I(\delta)$; under (1), (4) and (5), y_t and x_t are said to be cointegrated $CI(\delta, \gamma)$, for which a necessary condition is that y_t and x_t share the same order of integration, the latter terminology referring to the argument of $I(\cdot)$. The truncations on the right hand side in (1) ensure that the model is well-defined in the mean square sense, for example $\Delta^{-\delta} u_{2t}$ does not have finite variance when $\delta \geq 1/2$.

A feature referred to in the paragraph above is that we anticipate

$$Cov(u_{1t}, u_{2t}) \neq 0, \tag{6}$$

so that, viewing the first equation of (1) as a regression model, the regressor x_t is contemporaneously correlated with the cointegrating error $u_{1t}^{\#}$. The cointegrating vector mentioned above is the scalar ν in the bivariate setting (1), while we referred there to the case $\gamma = 0$, $\delta = 1$. The most dramatic contrast with this familiar $CI(1, 0)$ situation arises when

$$\delta < 1/2, \tag{7}$$

because the "simultaneous equation bias" inherent in (6) leads to inconsistency of the LSE due to the fact that x_t is asymptotically stationary and so its sum of squares does not asymptotically dominate that of u_{1t}. To overcome this problem, Robinson (1994) showed that a narrow-band frequency-domain least squares estimate (NBLSE) is consistent, due to the dominance near zero frequency of an $I(\gamma)$ spectral density by an $I(\delta)$ one. (He considered the purely stationary situation, where there is no truncation in (1), but our modification does not affect such basic asymptotic properties.) The same method was subsequently studied by Robinson and Marinucci (1997, 1999) in case

$$\delta > 1/2, \tag{8}$$

where there is trending nonstationarity. Here, the LSE is consistent, with convergence rate depending on the location of γ and δ in the non-negative quadrant, but the NBLSE still sometimes converges faster, and never converges slower, despite dropping high frequency information. Referring to a sequence m such that $m^{-1} + m/n \to 0$ as $n \to \infty$, the respective rates are: for $\gamma + \delta < 1$, $n^{2\delta-1}$ (LSE) and $n^{2\delta-1}(n/m)^{1-\gamma-\delta}$ (NBLSE); for $\gamma + \delta = 1$ but $\delta < 1$, $n^{2\delta-1}/\log n$ (LSE) and $n^{2\delta-1}/\log m$ (NBLSE); for $\gamma = 0$, $\delta = 1$, both estimates have rate n but the NBLSE enjoys less "second-order bias"; and for $\gamma + \delta > 1$, both have rate $n^{\delta-\gamma}$.

The question which then arises is whether these rates are optimal, by which we mean whether they match the rates achieved by the Gaussian maximum likelihood estimate (MLE) under suitable regularity conditions. They are optimal for the combination $\gamma+\delta > 1$, $\delta-\gamma > 1/2$, but otherwise not. In particular, the $n^{\delta-\gamma}$ rate is optimal for $\delta - \gamma > 1/2$ without the restriction $\gamma + \delta > 1$, and Robinson and Hualde (2000) have established it for estimates asymptotically equivalent to the MLE, allowing for consistent estimation of unknown γ and δ and a vector ψ of unknown parameters describing the autocovariance structure of u_t; these estimates of ν have mixed normal asymptotics, and a Wald test statistic with an asymptotic null χ^2 distribution, as established earlier in the $CI(1,0)$ case by Phillips (1991), Johansen (1991). Indeed, Robinson and Hualde (2000) found the limit distribution unaffected by the question of whether ψ, γ and δ are known or unknown. For related results on fractionally cointegrated models, with somewhat different estimates, settings and conditions, see Dolado and Marmol (1997), Jeganathan (1999, 2001), Kim and Phillips (2000). Testing for fractional cointegration is considered by Marinucci and Robinson (2001), Robinson and Yajima (2000).

The present paper focuses on the case

$$\beta \overset{def}{=} \delta - \gamma < 1/2, \tag{9}$$

where a substantially different asymptotic inferential theory prevails, impacting also on the question of how δ and γ should be estimated. Under (9), since $\Delta^\gamma y_t$ and $\Delta^\gamma x_t$ are $I(\beta)$, they are asymptotically stationary, and so, intuitively, one anticipates the existence of $n^{1/2}$-consistent and asymptotically normal estimates of ν; the LSE and NBLSE converge slower than this owing to the dominance of bias due to (6). Note that (9) excludes the traditional $CI(1,0)$ case and so might be thought of as less plausible than $\beta \geq 1/2$. However, the vast bulk of the cointegration literature has focused only on the $CI(1,0)$ possibility and there has been little study of fractional possibilities, or even the testing of the unit root hypothesis on y_t, x_t against fractional alternatives, as distinct from stationary autoregressive (AR) ones. In fact, the fractional cointegration analysis by Robinson and Marinucci (1997) of two of the bivariate series originally analysed by Engle and Granger (1987) (namely M1/nominal GNP and M3/nominal GNP) and one analysed by Campbell and Shiller (1987) (stock prices/dividends) in the $CI(1,0)$ context was suggestive of (9). Moreover, we cover not only $\delta \geq 1/2$, but also the asymptotically stationary case $\delta < 1/2$, which may be relevant for many financial time series. Note that here, the NBLSE of Robinson (1994) is only $m^{1/2}$-consistent for m increasing slower than n (indeed the optimal minimum-mean-squared error rate is $n^{2/5}$), so that we again achieve an improvement. We refer to (9) as "weak fractional cointegration", since the memory reduction achievable is small relative to the $CI(1,0)$ case, or other cases in which $\beta \geq 1/2$.

We are principally concerned with estimation of ν. If γ and δ are known, while u_t is known to be white noise with unknown variance-covariance matrix Ω, then the MLE of ν is given in closed form, and may be computed by means of an added-variable least squares regression, as pursued in the following section, which also extends to vector autoregressive (VAR) u_t, of known degree, but with unknown AR coefficients, when our estimate of ν is no longer as efficient as the MLE but has the same, $\sqrt{n}$, rate of convergence, under (9). When γ and/or δ are unknown, and u_t has parametric autocorrelation (such as following a VAR), then it seems that the Gaussian MLE of all the unknowns is again $\sqrt{n}$-consistent and asymptotically normal, but with limit covariance matrix that is not block-diagonal, so that in particular the asymptotic variance of the estimate of ν differs from that when γ and δ are known. If $\delta < 1/2$, a priori, conveying the implication that δ and γ are both estimated by optimizing over subsets of the intersection of (4) and (7), then the consistency and asymptotic distribution theory would largely follow the lines of authors such as Fox and Taqqu (1986) and Hosoya (1997), who were the first to develop such theory for standard scalar and vector long memory time series models respectively, the most notable difference perhaps being the fact that in our setting x_t and y_t would be only asymptotically stationary. If the possibility that $\delta \geq 1/2$ is admitted, and possibly $\gamma \geq 1/2$ also, then the situation is more delicate, as discussed in Section 4.

The preceding discussion makes it apparent that when γ and δ are unknown the issue of how they are estimated is of greater significance when $\beta < 1/2$ than when $\beta > 1/2$. It is indeed essential here (due to the correlation between x_t and u_{1t}) that they be estimated $\sqrt{n}$-consistently in order for ν to then be estimated $\sqrt{n}$ -consistently, so that simple closed-form semiparametric methods such as log periodogram regression will not suffice. Closed-form $\sqrt{n}$-consistent estimates of integration orders are available (see Kashyap and Eom, 1988, Moulines and Soulier, 1999), but these do not cover our bivariate situation and VAR u_t, and also entail logging the periodogram, which raises technical difficulties not present in estimates based on quadratic forms, such as the MLE. In our setting some degree of numerical optimization seems inevitable. Since this is likely to entail an initial search of the parameter space to locate the vicinity of a global optimum, it is desirable if the computations can be arranged so that only univariate optimizations are involved. Even after concentrating out parameters, when both γ and δ are unknown the Gaussian MLE requires a bivariate optimization under white noise u_t, and at least a trivariate optimization when u_t is VAR. We propose $\sqrt{n}$-consistent and asymptotically normal estimates that require only univariate optimizations.

The basic structure of the estimates of ν is described in the following section. Section 3 provides asymptotic theory in case γ and δ are known. Section 4 considers estimation of γ and δ and the effect on estimating ν. Section 5 contains Monte Carlo evidence of finite sample behaviour, and Section 6 several empirical applications.

2. Estimation of ν

We can write (1) as

$$z_t(\gamma,\delta) = \zeta x_t(\gamma)\nu + u_t^{\#}, \tag{10}$$

where we introduce the notation

$$v_t(c) = \Delta^c v_t^{\#}, \tag{11}$$

for a generic sequence v_t, and define

$$z_t(c,d) = (y_t(c), x_t(d))', \quad \zeta = (1,0)'. \tag{12}$$

We take u_t to be generated by the VAR

$$u_t = \sum_{j=1}^{p} B_j u_{t-j} + \varepsilon_t, \tag{13}$$

where all zeros of $\det\{I_2 - \sum_{j=1}^{p} B_j z^j\}$ lie outside the unit circle, the B_j being 2×2 matrices, and I_r the $r\times r$ identity matrix, while ε_t is a bivariate sequence, uncorrelated and homoscedastic over t, with mean zero and covariance matrix Ω. We take (13) to mean white noise u_t when $p = 0$.

From (10) and (13) we have

$$z_t(\gamma,\delta) - \sum_{j=1}^{p} B_j z_{t-j}(\gamma,\delta) = \nu\left\{\zeta x_t(\gamma) - \sum_{j=1}^{p} B_j \zeta x_{t-j}(\gamma)\right\} + \varepsilon_t^{+}, \quad t \geq 1, \tag{14}$$

where

$$\begin{aligned}
\varepsilon_1^{+} &= u_1, \\
\varepsilon_t^{+} &= u_t - \sum_{j=1}^{t-1} B_j u_{t-j}, \quad t = 2, ..., p, \\
\varepsilon_t^{+} &= \varepsilon_t, \quad t > p.
\end{aligned} \tag{15}$$

Denote by B_{ij} the ith row of B_j. Writing ε_{it} for the ith element of ε_t, for $t > p$, the second equation of (14) can be written as

$$x_t(\delta) - \sum_{j=1}^{p} B_{2j} z_{t-j}(\gamma,\delta) = -\nu \sum_{j=1}^{p} B_{2j}\zeta x_{t-j}(\gamma) + \varepsilon_{2t}, \tag{16}$$

whence the first equation can be written as

$$y_t(\gamma) = \nu x_t(\gamma) + \rho x_t(\delta) + \sum_{j=1}^{p}\left(B_{1j} - \rho B_{2j}\right) z_{t-j}(\gamma,\delta) - \nu \sum_{j=1}^{p}\left(B_{1j} - \rho B_{2j}\right)\zeta x_{t-j}(\gamma) + \varepsilon_{1.2,t}, \tag{17}$$

where $\varepsilon_{1.2,t} = \varepsilon_{1t} - \rho\varepsilon_{2t}$, $\rho = E(\varepsilon_{1t}\varepsilon_{2t})/E(\varepsilon_{2t}^2)$; (17) is a form of error-correction representation.

We wish to cater for the possibility of prior zero restrictions on the B_j which serve to eliminate some $y_{t-j}(\gamma)$, $x_{t-j}(\gamma)$, $x_{t-j}(\delta)$, as this will improve efficiency. Thus we introduce a $q \times (3p+2)$ matrix, which is I_{3p+2} when there are no such restrictions, but for $q < 3p+2$, Q is formed by dropping rows corresponding to the restrictions. Thus we can write (17) as

$$y_t(\gamma) = \theta' Q Z_t(\gamma,\delta) + \varepsilon_{1.2,t}, \tag{18}$$

where

$$Z_t(c,d) = \left(x_t(c), x_t(d), w'_{t-1}(c,d), ..., w'_{t-p}(c,d)\right)', \tag{19}$$

$$w_t(c,d) = \left(x_t(c), x_t(d), y_t(c)\right)'. \tag{20}$$

Since $E(\varepsilon_{1.2,t} Z_t(\gamma,\delta)) = 0$, we consider the (possibly constrained) least squares estimate

$$\widehat{\theta}(c,d) = G(c,d)^{-1} g(c,d), \tag{21}$$

taking $(c,d) = (\gamma,\delta)$, $(\gamma,\widetilde{\delta})$, $(\widetilde{\gamma},\delta)$ or $(\widetilde{\gamma},\widetilde{\delta})$, depending on whether γ and/or δ are known or estimated by $\widetilde{\gamma},\widetilde{\delta}$, and

$$G(c,d) = Q\frac{1}{n}\sum_{t=p+1}^{n} Z_t(c,d) Z'_t(c,d) Q', \quad g(c,d) = Q\frac{1}{n}\sum_{t=p+1}^{n} Z_t(c,d) y_t(c). \tag{22}$$

For example, in case $p=1$, if u_{1t} is white noise while u_{2t} is AR(1), then $q=3$ and (17) becomes

$$y_t(\gamma) = \nu x_t(\gamma) + \rho x_t(\delta) - \rho B_{221} x_{t-1}(\delta) + \varepsilon_{1.2,t}, \tag{23}$$

where B_{22j} is the second element of B_{2j}. Notice that ν, ρ and B_{221} are all identified in (23), but it is apparent from comparison of (17) with (18) that in general, while ν and ρ are expected to be identified, only some elements of the B_j are. However, we are treating the B_j as nuisance parameters, indeed it is principally ν that is of interest, so we stress

$$\widehat{\nu}(c,d) = 1' G(c,d)^{-1} g(c,d), \tag{24}$$

where $1 = (1,0,...,0)'$.

The representation (17) is of error-correction type and in case $p=0$, $\widehat{\nu}(\gamma,\delta)$ actually provides the Gaussian MLE of ν, given knowledge of γ,δ but lack of knowledge of Ω. For $p \geq 1$, it is less efficient than the MLE for this case, but still $n^{1/2}$-consistent and computationally considerably simpler. Notice that over-specification of p results in a further efficiency loss, but under-specification of p produces inconsistency. In moderate sample sizes, a modest choice of p, even $p=1$, might thus be a wise precaution. On the other hand, one could also regard (13) as approximating a more general infinite AR process with nonparametric $I(0)$ autocorrelation.

3. Asymptotic Theory with Known γ,δ

The present section establishes the $n^{1/2}$-consistency and asymptotic normality of $\widehat{\theta}(\gamma,\delta)$, and hence of $\widehat{\nu}(\gamma,\delta)$. We assume in addition to the description of (13) that the ε_t are stationary and ergodic with finite fourth moment, satisfying also

$$E\left(\varepsilon_t \middle| \mathcal{F}_{t-1}\right) = 0, \quad E\left(\varepsilon_t\varepsilon'_t \middle| \mathcal{F}_{t-1}\right) = \Omega \tag{25}$$

almost surely, where $\mathcal{F}_t$ is the σ-field of events generated by ε_s, $s \leq t$, and also assume that conditional (on $\mathcal{F}_{t-1}$) third and fourth moments and cross-moments of elements of ε_t equal the corresponding unconditional moments. Thus, the ε_t essentially behave like an iid sequence up to 4th moments. Now, noting from (1) that

$$x_t(\gamma) = \sum_{j=0}^{t-1} a_j(\beta) u_{2,t-j}, \quad t > 0; \quad = 0, \quad t \leq 0, \tag{26}$$

define

$$\overline{x}_t(\gamma) = \sum_{j=\max(t,0)}^{\infty} a_j(\beta) u_{2,t-j}, \quad \widetilde{x}_t(\gamma) = x_t(\gamma) + \overline{x}_t(\gamma), \tag{27}$$

so that because of (9), $\widetilde{x}_t(\gamma)$, $t = 0, \pm 1, ...$, is a covariance stationary sequence. Likewise, so is

$$\widetilde{y}_t(\gamma) = \nu \widetilde{x}_t(\gamma) + u_{1t}, \tag{28}$$

as is u_{2t}. Now define

$$\widetilde{w}_t = (\widetilde{x}_t(\gamma), u_{2t}, \widetilde{y}_t(\gamma))', \quad \widetilde{Z}_t = \left(\widetilde{x}_t(\gamma), u_{2t}, \widetilde{w}'_{t-1}, ..., \widetilde{w}'_{t-p}\right)', \tag{29}$$

$$\Phi = E(\widetilde{Z}_t \widetilde{Z}'_t), \quad \Psi = E\left(\varepsilon^2_{1.2,t} \widetilde{Z}_t \widetilde{Z}'_t\right). \tag{30}$$

The proof of the following theorem is left to Appendix A.

Theorem 3.1 As $n \to \infty$

$$n^{1/2}\left\{\widehat{\theta}(\gamma,\delta) - \theta\right\} \to_d N\left(0, (Q\Phi Q')^{-1} Q\Psi Q' (Q\Phi Q')\right)^{-1}, \tag{31}$$

and the covariance matrix on the right hand side is consistently estimated by

$$G(\gamma,\delta)^{-1} K(\gamma,\delta) G(\gamma,\delta)^{-1}, \tag{32}$$

where

$$K(c,d) = Q \frac{1}{n} \sum_{t=p+1}^{n} \widehat{\varepsilon}^2_{1.2,t}(c,d) Z_t(c,d) Z'_t(c,d) Q', \tag{33}$$

in which

$$\widehat{\varepsilon}_{1.2,t}(c,d) = y_t(c) - \widehat{\theta}(c,d)' Q Z_t(c,d). \tag{34}$$

Remark 3.1 For $p \geq 1$, $\widehat{\nu}(\gamma,\delta)$ is inefficient relative to the Gaussian MLE. Overparameterization in the B_j results in further loss of efficiency in estimation of ν. Consider the case where, in the estimation, the B_j are taken to be diagonal, with also u_{1t} white noise and u_{2t} $AR(p)$, to extend (23). Then, if in fact u_{2t} is also white noise the limiting variance of $n^{1/2}\{\widehat{\nu}(\gamma,\delta) - \nu\}$ is

$$\omega^2_{1.2} / \sum_{j=p+1}^{\infty} a^2_j(\beta), \tag{35}$$

where $\omega^2_{1.2} = E(\varepsilon^2_{1.2,t})$; (35) is increasing in p. As a simpler alternative to (33), (34), we can consistently estimate (35) by

$$\widehat{\omega}^2_{1.2}(\gamma,\delta) \left(1' G(\gamma,\delta) 1\right)^{-1}, \tag{36}$$

where

$$\widehat{\omega}^2_{1.2}(\gamma,\delta) = \frac{1}{n} \sum_{t=p+1}^{n} \widehat{\varepsilon}^2_{1.2,t}(\gamma,\delta). \tag{37}$$

Note that (35) and (36) also apply in case $p = 0$ is correctly taken in the estimation, when $\widehat{\nu}(\gamma, \delta)$ is equivalent to the Gaussian MLE, and (35) becomes

$$\omega_{1.2}^2 / \left\{ \frac{2^{-4\beta}}{\pi} B\left(1/2 - \beta, 1/2 - \beta\right) - 1 \right\}. \tag{38}$$

Note also that (35) and (38) do not depend on fourth cumulants of ε_t. However, if in fact u_t is not white noise, the limiting variance of $n^{1/2}\{\widehat{\nu}(\gamma, \delta) - \nu\}$, namely $1'(Q\Phi Q')^{-1}Q\Psi Q'(Q\Phi Q')^{-1}1$ (see (31)), in general depends on the fourth cumulant of $\varepsilon_{1.2,t}$, $\varepsilon_{1.2,t}$, ε_{2t} and ε_{2t}, though of course this is zero under Gaussianity.

Remark 3.2 On the other hand, under-parameterization of the B_j produces inconsistency of $\widehat{\nu}(\gamma, \delta)$, as when u_t is actually $AR(p+1)$. In this connection, note that Robinson and Hualde (2000) considered the Gaussian MLE for $\beta > 1/2$ in case of a far more general parametric class than (13). We can view (13) more informally, as approximating an actual, unknown, time series model in the hope that bias is decreasing in p, a statement which can likely be justified in a rigorous way by allowing p to increase slowly with n. Our AR approach is computationally convenient, and is in a long tradition of macroeconometric estimation of linear simultaneous equations systems, as well as relating to Johansen's (1991) approach to $CI(1, 0)$ cointegration. In case of ARMA models, over-parameterization of both AR and MA orders can have more serious consequences than those discussed in Remark 3.1.

Remark 3.3 So long as $p \geq 1$ and some B_j are non-diagonal, the endogeneity property (6) holds even when Ω is diagonal, i.e. $\rho = 0$.

4. The Case of Unknown γ, δ

The main practical interest in fractional cointegration centres on the realistic situation in which γ and/or δ are unknown. We shall focus on the case where both γ and δ are unknown, as being the most difficult both computationally and theoretically.

First, suppose that u_t is correctly taken to be white noise, with unknown covariance matrix Ω satisfying (6). Considering the Gaussian log-likelihood, both Ω and ν can be concentrated out to leave an objective function of γ and δ. The resulting estimates of γ and δ can then be plugged into (24). As mentioned in Section 1, asymptotic theory under $\delta < 1/2$ is a relatively standard extension of that for Gaussian estimates in such models as stationary fractional ARIMAs. For fractional ARIMAs whose integration order is allowed to take nonstationary values, there has been difficulty with the consistency proof (an essential preliminary to limit distribution theory, because estimates are only implicitly defined). This is especially due to lack of uniformity of convergence of the objective function around admissible values 0.5 less than the true value of the integration order, as discussed by Velasco and Robinson (2000), who by means of tapering, and a different definition of fractional nonstationarity from ours, established $\sqrt{n}$-consistent and asymptotically normal frequency-domain estimation of integration orders and other parameters in quite general univariate models, while allowing the admissible set to be arbitrarily large. Tapering, however, inflates the variance, while time domain estimates conveniently exploit the simple white noise or VAR structure of u_t, and seem natural for our definition of nonstationarity, and are certainly justifiable if δ and γ are known to lie in intervals of length no greater than $1/2$, for example $(0, 1/2)$ or $(1/2, 1]$.

We propose estimates of γ, δ and ν that are $\sqrt{n}$-consistent and asymptotically normal and require two univariate nonlinear optimizations, in place of one bivariate one. Our procedure extends nicely to the VAR u_t case, where after cancelling out Ω and the B_j, the Gaussian MLE is a trivariate function; note that ν and the B_j are involved bilinearly as well as linearly in (14).

Pursuing the case of white noise u_t, i.e. $p = 0$ in (13), we can write the second equation of (1) as

$$x_t(\delta) = \varepsilon_{2t}, \quad t \geq 1. \tag{39}$$

It is proposed to estimate δ by

$$\widetilde{\delta}_0 = \arg\min_{d \in \mathcal{D}} S_0(d), \tag{40}$$

for a compact set $\mathcal{D}$ and

$$S_0(d) = \sum_{t=1}^{n} x_t^2(d). \tag{41}$$

Then, we estimate γ by

$$\widetilde{\gamma}_0 = \arg\min_{c \in \mathcal{C}} T_0(c), \tag{42}$$

for a compact set $\mathcal{C}$ (presumably a subset of $[0, \widetilde{\delta}]$) and

$$T_0(c) = \sum_{t=1}^{n} \left\{ y_t(c) - \widehat{\nu}(c, \widetilde{\delta}_0) x_t(c) - \widehat{\rho}(c, \widetilde{\delta}_0) x_t(\widetilde{\delta}_0) \right\}^2, \tag{43}$$

where $\widehat{\nu}(c, d)$ is given by (24), taking $p = 0$, and $\widehat{\rho}(c, d)$ is the second element of $\widehat{\theta}(c, d)$ in this case. Notice that the presence of c as argument in $y_t(c)$, and indeed of d in $x_t(d)$ of (41), presents no barrier to consistent estimation because, for example, $y_t(c)$ involves c only in the coefficients of lagged values $y_{t-1}, y_{t-2}, ...,$ not y_t.

In case of VAR u_t, we develop further the triangular structure of (1) by assuming

$$B_j \text{ is upper-triangular, } j = 1, ..., p. \tag{44}$$

This corresponds to a kind of causal structure, with y_t formed from $y_{t-1}, y_{t-2}, ...$ and $x_t, x_{t-1}, ...,$ but x_t being determined by

$$x_t(\delta) - \phi' R X_t(\delta) = \varepsilon_{2t}, \tag{45}$$

with

$$X_t(d) = (x_{t-1}(d), ..., x_{t-p}(d))', \tag{46}$$

and R an $r \times p$ matrix with $R = I_p$ in case $r = p$ but for $r < p$ R is formed by dropping specified rows from I_p, in case $B_{22j} = 0$ for some j. The prescription (45) includes the case of diagonal B_j, does not seem an excessive requirement given the allowance for non-diagonal Ω, and introduces an element of parsimony.

Define

$$\widehat{\phi}(d) = H(d)^{-1} h(d), \tag{47}$$

where

$$H(d) = R \frac{1}{n} \sum_{t=p+1}^{n} X_t(d) X_t'(d) R', \quad h(d) = R \frac{1}{n} \sum_{t=p+1}^{n} X_t(d) x_t(d). \tag{48}$$

First, estimate δ by

$$\widetilde{\delta}_p = \arg\min_{d \in \mathcal{D}} S_p(d), \tag{49}$$

where

$$S_p(d) = \sum_{t=p+1}^{n} \left\{ x_t(d) - \widehat{\phi}(d)' R X_t(d) \right\}^2. \tag{50}$$

Then, estimate γ by

$$\widetilde{\gamma}_p = \arg\min_{c \in \mathcal{C}} T_p(c), \tag{51}$$

where

$$T_p(c) = \sum_{t=p+1}^{n} \left\{ y_t(c) - \widehat{\theta}(c, \widetilde{\delta}_p)' Q Z_t(c, \widetilde{\delta}_p) \right\}^2. \tag{52}$$

As abbreviating notation, we denote throughout, for any $p \geq 0$, $\widetilde{\delta} = \widetilde{\delta}_p$, $\widetilde{\gamma} = \widetilde{\gamma}_p$. In the following theorem, we assume $\gamma \in \mathcal{C}$, $\delta \in \mathcal{D}$ and take the supports of $\mathcal{C}$ and $\mathcal{D}$ to be of width less than 0.5 to avoid a difficulty described earlier in this section. The proof is omitted as it is extremely complicated and lengthy, while not entailing any novel difficulty.

Theorem 4.1 As $n \to \infty$

$$n^{1/2} \begin{bmatrix} \widehat{\nu}(\widetilde{\gamma}, \widetilde{\delta}) - \nu \\ \widetilde{\gamma} - \gamma \\ \widetilde{\delta} - \delta \end{bmatrix} \to_d N\left(0, ABA'\right), \tag{53}$$

where A is a $3 \times (q+2)$ matrix and B is a $(q+2) \times (q+2)$ matrix, for which consistent estimates $\widehat{A}$ and $\widehat{B}$ are presented in Appendix B.

Remark 4.1 Analytic formulae, in either the time or frequency domain, for A and B are excessively complicated, and thus omitted. Note that the estimate $\widehat{A}\widehat{B}\widehat{A}'$ provided by Appendix B is guaranteed non-negative definite.

Remark 4.2 As well as being useful in inference on ν, the theorem could also be applied in inference on γ and δ, for example to set a confidence interval for β which could be useful in judging the suitability of the weak cointegration specification (9).

Remark 4.3 On the other hand, our estimation procedure, though not our asymptotic theory, can also be used when $\beta > 1/2$, though alternative, possibly computationally more convenient, methods, are available here.

Remark 4.4 One approach, suggested in Robinson and Hualde (2000) when $\beta > 1/2$, is the use of residuals from LSE or NBLSE estimates of ν in the estimation of γ. However, these are always less-than-$n^{1/2}$-consistent under (9), and so it appears that the resulting estimates of γ will not achieve the essential $n^{1/2}$-consistency needed to provide an $n^{1/2}$-consistent estimate of ν.

Remark 4.5 Even when u_t is white noise, $\widehat{\nu}(\widetilde{\gamma}, \widetilde{\delta})$, $\widetilde{\delta}$ and $\widetilde{\gamma}$ are inefficient relative to the Gaussian MLE; intuitively, this is due to the estimation of δ from only the second equation of (1) (i.e. (40)), whereas the first equation also contains relative information. However, the estimates can be updated to efficiency by a single Newton step.

5. Monte Carlo Evidence

With the main aim of investigating the performance in finite samples of the estimates of ν proposed in the paper and associated rules of inference, and making comparisons with the simplest estimate, the LSE, a Monte Carlo experiment was carried out. In data generation from (1), (13), we took $p = 1$ throughout, with

$$B_1 = diag\left\{b_1, b_2\right\}, \tag{54}$$

where each of the b_i was allowed to take each of the values 0, 0.5, 0.9. The case $b_1 = b_2 = 0$ actually corresponds to $p = 0$ in (13), where u_t is a white noise vector. Likewise, $b_1 = 0$, $b_2 \neq 0$ corresponds to (23). We have employed in (54) abbreviating notation compared to (23), so $b_2 = B_{221}$. The ε_t in (13) were generated as Gaussian with $E(\varepsilon_{1t}^2) = E(\varepsilon_{2t}^2) = 1$ and $E(\varepsilon_{1t}\varepsilon_{2t}) = \rho$, taking values -0.5, 0, 0.5, 0.75, via the g05ezf routine of the Fortran NAG library. We varied ρ in order to assess possible "simultaneous equation bias", x_t and u_{1t} being orthogonal only when $\rho = 0$. We employed four (γ, δ) combinations:

$$(\gamma, \delta) = (0, 0.4)\,,\ (0.2, 0.4)\,,\ (0.4, 0.8)\,,\ (0.7, 1)\,, \tag{55}$$

for all of which $\beta < 1/2$. Notice that variances of all estimates, both in finite samples and asymptotically, will inevitably vary across parameter values. For example, because the $E(\varepsilon_{it}^2)$ are fixed throughout, $E(\varepsilon_{1.2,t}^2)$ will decrease in $|\rho|$, while $E(u_{it}^2)$ will increase in b_i. Finite sample biases of our estimates will doubtless also be affected by such variation, though in a more subtle manner. We took $\nu = 1$.

For each combination of parameter values, 1000 series of $\{y_t, x_t\}$ of lengths $n = 64, 128, 256$ were generated. Fractional series were generated as in (26), using $a_0(\alpha) = 1$, $a_{j+1}(\alpha) = ((j+\alpha)/(j+1))a_j(\alpha)$, $j \geq 1$, for $\alpha > 0$. For each series, we computed estimates of the following three types:

(i) The LSE,

$$\overline{\nu}_0 = \sum_{t=1}^{n} x_t y_t / \sum_{t=1}^{n} x_t^2. \tag{56}$$

(ii) The Infeasible estimate $\overline{\nu}_I = \widehat{\nu}(\gamma, \delta)$ based on correct specification and misspecification and/or over-specification.

(iii) The Feasible estimate $\overline{\nu}_F = \widehat{\nu}(\widetilde{\gamma}, \widetilde{\delta})$ based on correct specification and misspecification and/or over-specification.

By "correct specification" we mean that all prior zero restrictions on B_1 in (54), including the non-diagonal ones and any diagonal ones, are incorporated in the estimation, but not equality restrictions. By "mis-specification" we mean that for $b_1 \neq 0$ and $b_2 \neq 0$ we took $Z_t(c, d) = (x_t(c), x_t(d))'$. By "over-specification" we mean that for $b_1 = b_2 = 0$ we took $Z_t(c, d) = \left(x_t(c), x_t(d), w_{t-1}'(c, d)\right)'$. Of course, knowledge of $\rho = 0$ was never used. Table I records the convergence rates of the LSE and, under the heading "optimal", of $\overline{\nu}_I$, $\overline{\nu}_F$.

TABLE I
CONVERGENCE RATES

(γ, δ)	$(0, 0.4)$	$(0.2, 0.4)$	$(0.4, 0.8)$	$(0.7, 1)$
Optimal	$n^{.5}$	$n^{.5}$	$n^{.5}$	$n^{.5}$
LSE, $\rho \neq 0$	inconsistent	inconsistent	$n^{.4}$	$n^{.3}$
LSE, $\rho = 0$	$n^{.5}$	$n^{.5}$	$n^{.4}$	$n^{.3}$

We describe how $\widetilde{\delta}$ and $\widetilde{\gamma}$ in $\overline{\nu}_F$ were computed. In estimating δ, we fixed $\mathcal{D} = [\widehat{\delta} - 0.15, \widehat{\delta} + 0.15]$ in (49), where $\widehat{\delta}$ is the version of the log periodogram estimate of Geweke and Porter-Hudak (1983) proposed by Robinson (1995a), applied to the series x_t without pooling or trimming, based on bandwidths $m = 20, 30, 60$, corresponding to $n = 64, 128, 256$, respectively in case u_{2t} is assumed in the estimation to be white noise, and on $m = 10, 15, 30$, corresponding to $n = 64, 128, 256$, in case u_{2t} is assumed in the estimation to be AR(1). In all cases, $\mathcal{D}$ contains the asymptotic 95% confidence interval $[\widehat{\delta} - 1.96 s.e.(\widehat{\delta}), \widehat{\delta} + 1.96 s.e.(\widehat{\delta})]$, where $s.e.(\widehat{\delta}) = \pi/\sqrt{24m}$ is the asymptotic standard error of $\widehat{\delta}$ (Robinson, 1995a). In estimating γ, we fixed $\mathcal{C} = [\widetilde{\delta} - 0.50, \widetilde{\delta} - 0.05]$ in (51). The lower bound corresponds to the assumption $\beta < 1/2$. The upper bound seems reasonable since a very small (less than 0.05) β is unlikely to be detectable, indeed there is then near loss of identificability and very poor behaviour of estimates of ν.

Tables 2-7 report Monte Carlo bias (defined as the estimate minus the true value) of $\overline{\nu}_0$, $\overline{\nu}_I$ and $\overline{\nu}_F$, each table referring to a particular (b_1, b_2) combination with either correct specification, mis-specification or over-specification. Only some of the combinations covered in the experiment are included, in order to conserve in space. Generally, $\overline{\nu}_I$ performs best, followed by $\overline{\nu}_F$, with $\overline{\nu}_0$ worst.

We discuss first the cases of correct specification (Tables 2-5). The overall ordering is found in the full white noise case $b_1 = b_2 = 0$ (Table 2), and in the AR case (Tables 3-5) when $\rho \neq 0$, but not when $\rho = 0$ with $b_1 = b_2 \neq 0$, where $\overline{\nu}_0$ is best. For $b_1 = b_2 = 0.9$, $(\gamma, \delta) = (0.7, 1)$ and small n, $\overline{\nu}_0$ usually beats $\overline{\nu}_F$ even when $\rho \neq 0$ (Table 3). For $b_1 = 0$, $b_2 \neq 0$ (Table 4), we are close to the white noise outcome, but when $b_1 \neq 0$, $b_2 = 0$ the bias of $\overline{\nu}_0$ decays very slowly, and is unacceptably large when $b_1 = 0.9$ (Table 5). Focusing now more on variation across (γ, δ), the bias of $\overline{\nu}_I$ decreases in β, as is the case for $\overline{\nu}_F$ when $b_1 = b_2 = 0$. With AR structure, the worst performance of $\overline{\nu}_F$ is generally found for $(\gamma, \delta) = (0.2, 0.4)$ or $(0.7, 1)$. As for $\overline{\nu}_0$, bias varies with collective memory $\gamma + \delta$ when $\rho = 0$, but when $\rho \neq 0$, $(0, 0.4)$ and $(0.2, 0.4)$ are the worst cases, unsurprisingly in view of the LSE's inconsistency here. Generally, $\overline{\nu}_F$ works best under $(0.4, 0.8)$. With respect to variation in ρ, overall, the bias shares the sign of ρ in case of $\overline{\nu}_0$, $\overline{\nu}_I$, but is opposite in case of $\overline{\nu}_F$. $\overline{\nu}_I$ is relatively insensitive to ρ, though for $b_1 = 0.9$, $b_2 = 0$ (Table 5), bias increases in $|\rho|$, as is the case for $\overline{\nu}_0$, but no clear pattern can be found in the results for $\overline{\nu}_F$, though there is evidence of increase in bias with $|\rho|$. Looking at variation across (b_1, b_2), AR structure tends to reduce bias in $\overline{\nu}_0$ but increase it, and possibly change its sign, in $\overline{\nu}_I$. For $\overline{\nu}_F$, the worst performances occur when $b_1 \neq 0$, but even here bias decays rapidly as n increases, as it does also for $\overline{\nu}_I$.

Mis-specification (Table 6) has surprisingly little effect on $\overline{\nu}_I$, but seriously damages $\overline{\nu}_F$, especially when β is small, $(0.7, 1)$ being clearly the worst case, though when $\beta = 0.4$, bias decreases with n. As anticipated, over-specification (Table 7) makes little difference to $\overline{\nu}_I$, $\overline{\nu}_F$, which do much better than $\overline{\nu}_0$.

Tables 8-11 contain Monte Carlo standard deviations for only a subset of the combinations for which bias results were displayed. As noted before, variability is considerably affected by parameter values, and the relative performance of $\overline{\nu}_0$, $\overline{\nu}_I$ and $\overline{\nu}_F$ can be illustrated by focusing on only few cases. In fact, $\overline{\nu}_0$ was superior to $\overline{\nu}_I$ for most of the combinations, including those not displayed, with $\overline{\nu}_F$ a poor third. With correct specification, this was most notably the case for small n and $b_1 = b_2 \neq 0$ (Table 9), in part due to the proliferation in regressors, five in $\overline{\nu}_I$ and $\overline{\nu}_F$ versus one in $\overline{\nu}_0$, with variability in $\widetilde{\delta}$ and $\widetilde{\gamma}$ considerably inflating standard deviations of $\overline{\nu}_F$ relatively to those of $\overline{\nu}_I$. Precision also increases with increasing n, and when one or both of the b_i is zero (see Tables 8 and 10), the performance of $\overline{\nu}_I$ and $\overline{\nu}_F$ improves relative to that of $\overline{\nu}_0$. On the other hand, with over-specification (Table 11), $\overline{\nu}_I$ and $\overline{\nu}_F$ unsurprisingly deteriorate further, and generally larger sample sizes will be required in order for their faster convergence rate to consistently deliver smaller standard deviations than $\overline{\nu}_0$. Nevertheless, it must be borne in mind that the paper's motivation is not to minimise variance but rather to achieve $n^{1/2}$-consistency and asymptotic normality in a fairly general context, which the LSE $\overline{\nu}_0$ does not provide.

We now go in to examine the usefulness of these limit distributional properties of $\overline{\nu}_I$ and $\overline{\nu}_F$ in finite-sample statistical inference, by examining the size of Wald tests. We computed

$$W_I = \frac{(\overline{\nu}_I - \nu)^2 n}{[G(\gamma,\delta)^{-1}K(\gamma,\delta)G(\gamma,\delta)^{-1}]_{(1)}},\ W_F = \frac{(\overline{\nu}_F - \nu)^2 n}{\left[\widehat{A}\widehat{B}\widehat{A}'\right]_{(1)}}, \quad (57)$$

where $[\cdot]_{(i)}$ denotes ith diagonal element. Empirical sizes, with respect to nominal sizes $\alpha = 0.05$ and 0.1, again across 1000 replications, are reported in Tables 12-17, for each of the (b_1, b_2) for which biases were given.

With correct specification, even for $b_1 = b_2 = 0$ (Table 12), the sizes of the infeasible statistic W_I are somewhat too large, and autocorrelation in u_t exarcebates this, with the case $b_1 \neq 0$, $b_2 = 0$ again worse than $b_1 = 0$, $b_2 \neq 0$, but not necessarily worse than $b_1 = b_2 \neq 0$ (Tables 13-15). Results for $\alpha = 0.1$ are clearly better than for $\alpha = 0.05$. Overall, there is improvement as n increases, and even for small n, the performance of W_I seems quite satisfactory. Predictably, mis-specification (Table 16) plays havoc, producing sizes that are unacceptably high, especially for $\alpha = 0.05$. With over-specification, performance is again good, though we would not expect high power.

For the feasible statistic W_F, with correct specification and no autocorrelation in u_t (Table 12), sizes are worse than for W_I, with less evidence of settling down as n increases and varying more across parameter values, sometimes actually being less than the nominal values. Indeed, with autocorrelation (Tables 13-15), sizes are emphatically too small and mostly further from the nominal values than the corresponding W_I are in the opposite direction, though this is by no means always the case, and for $n = 64$ and $\alpha = 0.05$ the results are extraordinarily good. However, we would not wish to draw over-optimistic general conclusions here, and certainly not from Table 16, where the mis-specification so evident in the results for W_I can barely be seen in those for W_F (though the superiority of W_F is even more dramatic when $b_1 = b_2 = 0.9$, for which results are not reported). With over-specification (Table 17), W_F mostly beats W_I, especially when $\alpha = 0.05$. It is possible that the performance of W_F relative to W_I is not accidental because W_I has an asymptotic formula in the denominator. Certainly, our overall experience with W_F is quite encouraging.

While we have stressed estimation of ν, estimates of δ and γ would also be of interest in any empirical analysis of fractional cointegration, and so we also give some space to the performance of $\widetilde{\delta}$ and $\widetilde{\gamma}$, and of Wald tests for δ and γ based on Theorem 4.1.

Tables 18 and 19 report Monte Carlo bias and standard deviation for $\widetilde{\delta}$ for the same values of δ (0.4, 0.8, 1), b_2 (0, 0.5, 0.9) and n (64, 128, 256) as before, again based on 1000 replications. However, we fix $\rho = 0.5$ here, using the same estimates of $\widetilde{\delta}$ computed in this case for the feasible estimates $\overline{\nu}_F$ and Wald statistics W_F discussed previously. We report results for minimization of both $S_0(d)$ and $S_1(d)$ (see (41), (50)), so that $S_0(d)$ with $b_2 = 0$ and $S_1(d)$ with $b_2 \neq 0$ both correspond to correct specification, $S_1(d)$ with $b_2 = 0$ to over-specification, and $S_0(d)$ with $b_2 \neq 0$ to mis-specification.

The biases based on $S_0(d)$ and $S_1(d)$ with $b_2 = 0$ increase somewhat with δ, but look satisfactory even for $n = 64$, and are decreasing in n. For $S_1(d)$ with $b_2 = 0.5$, there is some deterioration, but nevertheless performance is still acceptable, but for $b_2 = 0.9$, the results are very poor, even for $n = 256$, though this is not too surprising in view of the difficulties often caused by a near-unit root. Unsurprisingly, there is severe bias, increasing with b_2, when $S_0(d)$ is used with $b_2 \neq 0$. Standard deviations in the correctly specified and over-specified cases are pretty stable over δ, but, as expected, worse in the latter case.

Tables 20 and 21 report Monte Carlo sizes of Wald statistics for δ

$$W_\delta = \frac{(\widetilde{\delta} - \delta)^2 n}{\left[\widehat{A}\widehat{B}\widehat{A}'\right]_{(3)}}, \tag{58}$$

based on Theorem 4.1, with respect to the nominal sizes $\alpha = 0.05$, 0.1 respectively. As expected, under mis-specification they are far too large, and this is also the case using $S_1(d)$ with $b_2 = 0.9$. Otherwise, while still too large, they are not bad, and decrease in n, the ones for $\alpha = 0.1$ being best.

Tables 22-25 give corresponding results for $\widetilde{\gamma}$, with $b_1 = b_2 = b$ taking values 0, 0.5, 0.9, and for the four (γ, δ) combinations considered previously. Our estimation procedure being sequential, we consider two categories, $S_0(d)$ followed by $T_0(c)$ (43), and $S_1(d)$ followed by $T_1(c)$ (52), so that in the former case there is correct specification for $b = 0$ and mis-specification for $b \neq 0$, and in the latter, over-specification for $b = 0$ and correct specification for $b \neq 0$. The bias and standard deviation results of Tables 22 and 23 exhibit somewhat some variation across (γ, δ), but otherwise the qualitative conclusions for $\widetilde{\delta}$ still apply. With the Wald statistic

$$W_\gamma = \frac{(\widetilde{\gamma} - \gamma)^2 n}{\left[\widehat{A}\widehat{B}\widehat{A}'\right]_{(2)}}, \tag{59}$$

more variation in sizes is also found, in Tables 24 and 25, than for W_δ, some of the sizes being smaller than the nominal ones.

6. Empirical Examples

Using a methodology involving the LSE and NBLSE of ν, and semiparametric estimates of ν, Robinson and Marinucci (1997) found evidence that $\beta < 1/2$ in some of the bivariate macroeconomic series originally examined by Engle and Granger (1987), Campbell and Shiller (1987), who were investigating only the possibility of $CI\,(1,0)$ cointegration. This experience motivates application of our present approach to the same data. The main departure from the methodology of the previous section was an attempt at greater realism by determining p in (13) from the data, rather than assuming its value *a priori*. For this purpose, we need proxies for the u_{it}, which can only be obtained by operating on the observed y_t, x_t, series with preliminary estimates of ν, γ and δ. To estimate ν here we used the LSE $\overline{\nu}_0$, given by (56) (and computed by Robinson and Marinucci (1997)). To estimate γ and δ, we used semiparametric estimates (already computed by Robinson and Marinucci (1997), Marinucci and Robinson (2001)) in order to provide robustness against a range of short-memory specifications for u_t. Specifically, the estimates of γ and δ computed by these authors were of log periodogram (LP) and semiparametric Gaussian (SG) type (of the precise form considered by Robinson (1995a,b), using various bandwidths and based either on raw data/residuals or on first differenced ones followed by adding back 1. For asymptotic theory under stationarity we appeal to Robinson (1995a,b), and under nonstationarity, to Velasco (1999a,b). For preliminary estimates of γ, δ, ν, sample correlograms and partial correlograms were computed (to lag length 36) in order to identify, in the spirit of Box and Jenkins (1971), the AR orders of the u_{it}. For each data set, this was done for both the smallest and largest of the various univariate estimates based on the series x_t/residuals provided by Robinson and Marinucci (1997), Marinucci and Robinson (2001), and implications of both provided when the results could not be reconciled, recognizing the imprecision in semiparametric estimation.

We also took this opportunity to examine another question which in one form or another always arises with application of fractional models, and perhaps most acutely when nonstationary data are involved. This is the matter of truncation. When estimated innovations from a stationary fractional model are computed, the (infinite) AR representation has to be truncated because the data begins at time "1", not at time "$-\infty$". Now in our model (1) for nonstationary data, the truncation is actually inherent in the model, so strictly speaking there is no "error" associated with it. However, the model reflects the time when the data begins, and if we were to drop the first observation, say, and start the model off at the next one, the degree of filtering applied to all subsequent observations would change, and it is possible that this could have a marked effect, especially with nonstationary data, even though filtering is here applied after demeaning. To check for stability with respect to this phenomenon, we thus report computations based on the last $n' = n - j$ observations, for $j = 0, 1, ..., 10$.

We look first at Engle and Granger's (1987) quarterly consumption and income data, 1947Q1-1981Q2 ($n = 138$). They found evidence of $CI\,(1,0)$ cointegration, but did not investigate fractional possibilities. Marinucci and Robinson's (2001) analysis tends to support the notion of $\delta = 1$, but not of $\gamma = 0$, with positive estimates of γ that sometimes fall in the nonstationary region, thereby hinting that $\beta < 1/2$ is possible.

Taking y=consumption, x=income, the LSE of ν, from Robinson and Marinucci (1997), is 0.229. The two preliminary estimates of δ taken from Marinucci and Robinson (2001) were 0.89 (LP applied to first differences of x and adding back 1, with bandwidth 22) and 1.08 (SG applied to first differences of x and adding back 1, with bandwidth 40). In each case, the corresponding correlograms and partial correlograms suggested modelling u_{2t} as white noise. The preliminary estimates of γ were 0.19 (LP applied to raw residuals with bandwidth 22) and 0.87 (SG applied to first differenced residuals and adding back 1, with bandwidth 40). This large gap results in identification of an AR(1) u_{1t} in the first case, and white noise u_{1t} in the second. In view of these investigations, we carried out two distinct cointegration analyses, one with $p = 0$ in (13), the other with $p = 1$ in (13) with $B_1 = diag\,(b_1, 0)$.

In case u_{1t} and u_{2t} are both white noise, Table A reports values of the following statistics with n replaced by $n' = n - j$, $j = 0, ..., 10$: $\widehat{\nu} = \widehat{\nu}(\widetilde{\gamma}, \widetilde{\delta})$, $\widetilde{\delta}$, $\widetilde{\gamma}$, and their estimated standard errors $SE(\widehat{\nu})$, $SE(\widetilde{\delta})$, $SE(\widetilde{\gamma})$ from Theorem 4.1, $\widehat{\rho} = \widehat{\rho}(\widetilde{\gamma}, \widetilde{\delta})$, which is the estimated coefficient of $x_t(\widetilde{\delta})$ in (17) for $p = 0$ with $\widetilde{\gamma}$, $\widetilde{\delta}$, replacing γ, δ, and the correlation $Corr\,(\varepsilon_{1t}, \varepsilon_{2t})$ is estimated by

$$r = \widehat{\rho}(\widetilde{\gamma}, \widetilde{\delta})(\widehat{\sigma}_{22}/\widehat{\sigma}_{11})^{\frac{1}{2}}, \tag{60}$$

where

$$\widehat{\sigma}_{11} = n^{-1} {\sum_t}' \left(y_t(\widetilde{\gamma}) - \widehat{\nu}(\widetilde{\gamma}, \widetilde{\delta}) x_t(\widetilde{\gamma}) \right)^2, \; \widehat{\sigma}_{22} = n^{-1} {\sum_t}' x_t^2(\widetilde{\delta}), \tag{61}$$

with ${\sum_t}'$ meaning summation over the last n' observations.

As n' falls, $\widehat{\nu}$ and $\widetilde{\delta}$ tend to increase, and $\widetilde{\gamma}$ to decrease, but there is high stability for $n' \leq 133$, and generally the changes are insignificant relative to standard errors, $\widehat{\nu}$ for $n' = 128$ being one standard error larger than $\widehat{\nu}$ for $n' = 138$ (and also somewhat larger than the LSE). The estimates of δ and γ are certainly consistent with $\beta < 1/2$. More especially, exploiting the standard errors provided by our approach, the hypothesis that $\delta = 1$ seems rejectable against $\delta > 1$, but (though we do not report standard errors of $\widetilde{\beta} = \widetilde{\delta} - \widetilde{\gamma}$, which could be computed using Theorem 4.1) there is no evidence against $\beta < 1/2$. Substantial negative contemporaneous correlation between u_{1t} and u_{2t} is suggested. Note that dropping the first observation does not affect $\widetilde{\delta}$, as $x_1\,(d) = x_1$ for any d.

TABLE A

Consumption and Income: u_t white noise

n'	138	137	136	135	134	133	132	131	130	129	128
$\widehat{\nu}$	.223	.222	.251	.252	.251	.248	.247	.242	.243	.245	.246
$SE(\widehat{\nu})$	.027	.031	.024	.022	.023	.022	.023	.021	.022	.023	.023
$\widetilde{\delta}$	1.07	1.07	1.09	1.15	1.15	1.17	1.18	1.18	1.18	1.18	1.18
$SE(\widetilde{\delta})$	.028	.028	.059	.068	.073	.080	.083	.082	.083	.082	.084
$\widetilde{\gamma}$	.714	.745	.715	.692	.694	.696	.696	.685	.692	.694	.693
$SE(\widetilde{\gamma})$	.084	.092	.087	.087	.089	.090	.090	.089	.093	.093	.093
$\widehat{\rho}$	-.024	-.055	-.085	-.090	-.090	-.086	-.085	-.072	-.073	-.073	-.074
r	-.195	-.189	-.297	-.311	-.310	-.294	-.285	-.247	-.251	-.250	-.253

The analysis with u_{1t} AR(1) in Table B presents a very different picture. Here, we also report $\widehat{b}_1$ and $\widehat{\nu b_1}$, which are the estimated coefficients of $y_{t-1}(\widetilde{\gamma})$ and $-x_{t-1}(\widetilde{\gamma})$ in the regression (cf. (17)) used to compute $\widehat{\nu}$ and $\widehat{\rho}$, and $\widehat{\sigma}_{11}$ in r is now the sample average of the squared residuals from the regression of $y_t(\widetilde{\gamma}) - \widehat{\nu}(\widetilde{\gamma}, \widetilde{\delta}) x_t(\widetilde{\gamma})$ on $y_{t-1}(\widetilde{\gamma}) - \widehat{\nu}(\widetilde{\gamma}, \widetilde{\delta}) x_{t-1}(\widetilde{\gamma})$. In view of the AR(1) component, we effectively lose one observation, so n' goes from 127 to 137, the effect of then dropping the first observation being very striking, but the estimates subsequently exhibiting little variation across n'. As u_{2t} is still considered a white noise, the estimates of δ are identical to those of Table A, but estimates of γ are all now less than zero, although not significantly, Engle and Granger's (1987) $CI\,(1,0)$ conclusion now being supported. The AR component in u_{1t} clearly accounts for the bulk of the autocorrelation in cointegrating errors, resulting in the small estimates of γ, which are based on AR-transformed data. The MLE, which estimates γ simultaneously with b_1 and the other parameters, would allow AR and fractional features to compete more favourably, though, as discussed in the Introduction, it would require much heavier computation. Notice that $\widehat{\nu b_1}$ looks quite consistent with the values of $\widehat{\nu}$ and $\widehat{b}_1$, possibly providing some support for the present specification. Note also that the various $\widehat{\nu}$ are larger than before, but that, if indeed $\beta > 1/2$, their standard errors have to be interpreted with caution, as $\widehat{\nu}$ is then no longer asymptotically normal.

TABLE B

Consumption and Income: u_{1t} AR(1), u_{2t} white noise

n'	137	136	135	134	133	132	131	130	129	128	127
$\widehat{\nu}$	.163	.257	.264	.267	.263	.265	.258	.261	.262	.263	.262
$SE(\widehat{\nu})$	.179	.055	.054	.057	.053	.056	.051	.056	.055	.055	.054
$\widetilde{\delta}$	1.07	1.09	1.15	1.15	1.17	1.18	1.18	1.18	1.18	1.18	1.18
$SE(\widetilde{\delta})$	.028	.059	.068	.073	.080	.083	.082	.083	.082	.084	.084
$\widetilde{\gamma}$	-.101	-.167	-.183	-.184	-.184	-.179	-.193	-.180	-.184	-.189	-.186
$SE(\widetilde{\gamma})$	.234	.187	.181	.183	.185	.193	.180	.193	.192	.191	.192
$\widehat{b}_1$	.798	.843	.842	.839	.837	.832	.845	.842	.842	.842	.843
$\widehat{\nu b_1}$	.116	.221	.228	.230	.226	.226	.223	.225	.226	.227	.226
$\widehat{\rho}$	.009	-.088	-.102	-.104	-.102	-.105	-.093	-.096	-.094	-.095	-.094
r	.009	-.128	-.122	-.119	-.126	-.127	-.128	-.128	-.119	-.117	-.121

Engle and Granger (1987) found no evidence of $CI\,(1,0)$ cointegration between $\log M_1$ (y) and $\log GNP$ (x), on the basis of 90 quarterly observations, 1959Q1-1981Q2. Marinucci and Robinson's (1997) fractional analysis admitted the possibility of cointegration, with $\beta < 1/2$. In our preliminary analysis of autocorrelation in u_t, we took from their estimates of δ the values 1.22 (SG applied to first differences of x and adding back 1, using bandwidth 30) and 1.36 (LP applied to first differences of x and adding back 1, using bandwidth 22), and from their estimates of γ the values 0.76, 1.2, both LP estimates but applied respectively to raw residuals using bandwidth 22, and first differences of residuals and adding back 1, using bandwidth 16. Employing also the LSE of ν, 0.643, we found no evidence of autocorrelation in u_t, so proceeded to a cointegration analysis on the basis of $p = 0$ in (13). The results are reported in Table C. We found large variation across the largest n', but a good degree of stability is then achieved, with substantially larger values of $\widetilde{\delta}$ and $\widetilde{\gamma}$ (and of their standard errors). Clearly, $\widetilde{\delta}$ significantly exceeds 1, while $\widetilde{\gamma}$ does not, and the resulting $\widetilde{\beta} = \widetilde{\delta} - \widetilde{\gamma}$ are extremely close to the threshold value of 1/2. There is considerable negative correlation between u_{1t} and u_{2t}, and for the smaller n', $\widehat{\nu}$ is close to the LSE.

TABLE C

LogM1 and LogGNP: u_t white noise

n'	90	89	88	87	86	85	84	83	82	81	80
$\widehat{\nu}$	.704	.740	.578	.564	.608	.640	.638	.644	.643	.649	.658
$SE(\widehat{\nu})$	.077	.145	.040	.058	.058	.054	.054	.061	.061	.061	.061
$\widetilde{\delta}$	1.06	1.06	1.91	1.88	1.74	1.63	1.64	1.63	1.63	1.61	1.59
$SE(\widetilde{\delta})$	.057	.057	.025	.121	.117	.068	.083	.082	.086	.084	.076
$\widetilde{\gamma}$	.884	.928	1.12	1.16	1.11	1.09	1.09	1.11	1.10	1.10	1.09
$SE(\widetilde{\gamma})$	.108	.122	.121	.121	.131	.136	.138	.140	.140	.139	.139
$\widehat{\rho}$	-.134	-.222	-.261	-.268	-.315	-.352	-.350	-.379	-.376	-.391	-.408
r	-.839	-.543	-.402	-.413	-.455	-.475	-.473	-.507	-.504	-.515	-.522

Finally, we looked at the $n = 116$ annual observations, 1871-1986, on stock prices (y) and dividends (x), analysed by Campbell and Shiller (1987). Their findings with respect to $CI\,(1,0)$ cointegration were inconclusive, but Robinson and Marinucci's (1997) and Marinucci and Robinson's (2001) analyses again offered the possibility of cointegration with $\beta < 1/2$. The preliminary estimates of δ taken from Marinucci and Robinson (2001) were 0.86 and 0.95, being SG based on first differences of x and adding back 1, with bandwidths respectively 30 and 40. The preliminary estimates of γ were 0.57, 0.77, being LP on first differences of residuals and adding back one, with bandwidth 30, and SG on raw residuals with bandwidth 22, respectively. We also used the LSE of ν, 31. In this case, both γ estimates suggested white noise u_{1t}, while the δ estimates variously suggested white noise and AR(1) u_{2t}, but our subsequent fractional cointegration analysis produced $\widetilde{\gamma}$ and $\widetilde{\delta}$ that were too close to admit the likelihood of any cointegration. Thus, we report, in Table D, only the results with both u_{1t} and u_{2t} white noise. There is little variation with n', and strong support for the unit root hypothesis on δ, and, since $\widetilde{\gamma}$ is significantly larger than 1/2 at the 5% level, cointegration with $\beta < 1/2$ is certainly a possibility. We find that $\widehat{\nu}$ is somewhat larger than the LSE value, though not significantly so.

TABLE D

Stock Prices and Dividends: u_t white noise

n'	116	115	114	113	112	111	110	109	108	107	106
$\widehat{\nu}$	32.7	32.7	32.2	31.9	31.7	31.8	31.7	32.0	32.1	32.1	32.1
$SE(\widehat{\nu})$	7.56	7.64	7.80	7.83	7.81	7.93	7.91	7.99	8.02	7.99	8.01
$\widetilde{\delta}$	1.04	1.04	1.08	1.09	1.09	1.09	1.09	1.09	1.10	1.10	1.10
$SE(\widetilde{\delta})$	.077	.077	.090	.092	.092	.092	.093	.093	.095	.095	.095
$\widetilde{\gamma}$	.749	.751	.751	.752	.751	.752	.752	.751	.749	.749	.749
$SE(\widetilde{\gamma})$	.114	.116	.116	.117	.116	.117	.117	.116	.116	.116	.116
$\widehat{\rho}$	-8.97	-9.52	-9.13	-8.82	-8.56	-8.67	-8.54	-8.52	-.8.64	-8.59	-8.69
r	-.299	-.283	-.272	-.263	-.256	-.259	-.255	-.252	-.255	-.253	-.256

Appendix A: Proof of Theorem 3.1

We prove first that Φ is nonsingular, which ensures existence of the inverses in (31). Define

$$\Phi^{+} = E\left(\widetilde{Z}_t^{+}\widetilde{Z}_t^{+\prime}\right), \quad \widetilde{Z}_t^{+} = \left(\widetilde{w}_t', \widetilde{w}_{t-1}', ..., \widetilde{w}_{t-p}'\right)'. \tag{A.1}$$

It clearly suffices to show that Φ^+ is positive definite. Defining

$$\overline{\Phi}^+ = E\left(\overline{Z}_t \overline{Z}_t'\right), \quad \overline{Z}_t = \left(\overline{w}_t', \overline{w}_{t-1}', ..., \overline{w}_{t-p}'\right)', \tag{A.2}$$

for $\overline{w}_t = (\widetilde{x}_t(\gamma), u_{2t}, u_{1t})'$, from (28) it suffices to show that $\overline{\Phi}^+$ is positive definite, and similarly, defining

$$\overline{\Phi}^{++} = E\left(R\overline{Z}_t \overline{Z}_t' R'\right), \tag{A.3}$$

where R is a full rank $3(p+1) \times 3(p+1)$ matrix whose columns are orthonormal vectors such that

$$R\overline{Z}_t = [\overline{x}(\gamma)', \overline{u}_2', \overline{u}_1']', \tag{A.4}$$

where $\overline{x}(\gamma) = (\widetilde{x}_t(\gamma), ..., \widetilde{x}_{t-p}(\gamma))'$, $\overline{u}_2 = (u_{2t}, ..., u_{2,t-p})'$, $\overline{u}_1 = (u_{1t}, ..., u_{1,t-p})'$, it suffices to show that $\overline{\Phi}^{++}$ is positive definite. Define the vectors

$$e(\lambda) = \left(1, e^{i\lambda}, ..., e^{ip\lambda}\right)', \quad d(\lambda) = (1 - e^{i\lambda})^{-\beta} e(\lambda), \tag{A.5}$$

and the $3(p+1) \times 2$ matrix

$$E(\lambda) = \begin{bmatrix} 0' & 0' & e(\lambda)' \\ d(\lambda)' & e(\lambda)' & 0' \end{bmatrix}', \tag{A.6}$$

where $0'$ is here a $1 \times (p+1)$ vector of zeros. Define by $f(\lambda)$ the spectral density matrix of u_t, and note from positive finiteness of Ω and finiteness of the B_j that the smallest eigenvalue of the Hermitian matrix $f(\lambda)$ is bounded from below by a positive constant c, uniformly in λ. Then we can write

$$\overline{\Phi}^{++} = \int_{-\pi}^{\pi} E(\lambda) f(\lambda) E(-\lambda)' d\lambda, \tag{A.7}$$

which for some $c > 0$ exceeds

$$c\int_{-\pi}^{\pi} E(\lambda)E(-\lambda)'d\lambda = c \begin{bmatrix} A & B & 0 \\ B' & I_{p+1} & 0 \\ 0 & 0 & I_{p+1} \end{bmatrix} \tag{A.8}$$

by a non-negative definite matrix, where 0, A and B are $(p+1) \times (p+1)$ matrices, having (i,j)th elements 0, $\sum_{\ell=0}^{\infty} a_\ell a_{\ell+|i-j|}$ and $a_{j-i} 1(j \geq i)$ respectively, with $a_j = a_j(\beta)$. It thus suffices to show that $A - BB'$ is positive definite. But for a $(p+1) \times 1$ vector $\zeta = (\zeta_i)$,

$$\zeta'(A - BB')\zeta = \sum_{\ell=1}^{\infty} \left(a_\ell \zeta_{p+1} + ... + a_{\ell+p}\zeta_1\right)^2, \tag{A.9}$$

which is positive unless $\zeta = 0$ because $a_\ell / a_{\ell-1} = (\ell + \beta - 1)/\ell$ is strictly increasing in $\ell \geq 1$ for $\beta < 1$.

We now have to show that

$$\frac{1}{n} \sum{}' Z_t(\gamma, \delta) Z_t'(\gamma, \delta) \;\; \rightarrow \;\; {}_p\Phi, \tag{A.10}$$

$$n^{-1/2} \sum{}' Z_t(\gamma, \delta) \varepsilon_{1.2,t} \;\; \rightarrow \;\; {}_d N(0, \Psi), \tag{A.11}$$

writing $\sum' = \sum_{t=p+1}^{n}$. To prove (A.11), note first that it suffices to show

$$n^{-1/2} \sum{}' \widetilde{Z}_t \varepsilon_{1.2,t} \to_d N(0, \Psi), \tag{A.12}$$

because

$$\begin{aligned} E\left\| n^{-1/2} \sum{}' \left\{ Z_t(\gamma,\delta) - \widetilde{Z}_t \right\} \varepsilon_{1.2,t} \right\|^2 &\leq \frac{K}{n} \sum{}' E \left\| Z_t(\gamma,\delta) - \widetilde{Z}_t \right\|^2 \\ &\leq \frac{K}{n} \sum{}' \sum_{j=1}^{p} E\overline{x}_{t-j}^2(\gamma) \\ &\leq \frac{K}{n} \sum{}' \sum_{j=1}^{p} \int_{-\pi}^{\pi} \left| \sum_{s=t-j}^{\infty} a_s e^{-is\lambda} \right|^2 \|f(\lambda)\| \, d\lambda \\ &\leq \frac{K}{n} \sum_{t=1}^{n} \sum_{s=t}^{\infty} a_s^2 \to 0, \end{aligned} \tag{A.13}$$

as $n \to \infty$, by the Toeplitz lemma, the last inequality following because $f(\lambda)$ is bounded due to the assumption on the B_ℓ. Write $\widetilde{Z}_t = Z_{at} + Z_{bt}$, where the first two elements of Z_{at}, and the last $3p$ elements of Z_{bt}, equal corresponding ones of $\widetilde{Z}_t$. Thus Z_{bt} is $\mathcal{F}_{t-1}$-measurable and

$$E\left(\varepsilon_{1.2,t} \widetilde{Z}_t) \left| \mathcal{F}_{t-1} \right. \right) = E\left(\varepsilon_{1.2,t} Z_{at} \right) + Z_{bt} E\left(\varepsilon_{1.2,t} \left| \mathcal{F}_{t-1} \right. \right) = 0, \ a.s. \tag{A.14}$$

Further,

$$\begin{aligned} E\left(\varepsilon_{1.2,t}^2 \widetilde{Z}_t \widetilde{Z}_t^{\prime} \left| \mathcal{F}_{t-1} \right. \right) &= E\left(\varepsilon_{1.2,t}^2 Z_{at} Z_{at}' \right) + E\left(\varepsilon_{1.2,t}^2 Z_{at} \right) Z_{bt}' \\ &\quad + Z_{bt} E\left(\varepsilon_{1.2,t}^2 Z_{at}' \right) + E(\varepsilon_{1.2,t}^2) Z_{bt} Z_{bt}', \ a.s., \end{aligned} \tag{A.15}$$

and so

$$\frac{1}{n} \sum{}' \left[E\left\{ \varepsilon_{1.2,t}^2 \widetilde{Z}_t \widetilde{Z}_t' \left| \mathcal{F}_{t-1} \right. \right\} - E\left\{ \varepsilon_{1.2,t}^2 \widetilde{Z}_t \widetilde{Z}_t' \right\} \right] \to_p 0, \tag{A.16}$$

because Z_{bt} and $Z_{bt} Z_{bt}' - E(Z_{bt} Z_{bt}')$ are stationary and ergodic with zero means. Since (A.15) has expectation Ψ, (A.12) then follows from the Cramer-Wold device and Theorem 1 of Brown (1971), noting that the Lindeberg condition in the latter reference is trivially satisfied because $\varepsilon_{1.2,t} \widetilde{Z}_t$ is stationary with finite variance. Thus (A.11) is proved. The proof of (A.10) follows from (A.13) and elementary inequalities. This concludes the proof of (31). The proof of the final statement of the theorem is omitted as it is standard given (31) and its proof.

Appendix B: Definition of $\hat{A}$ and $\hat{B}$

For brevity we write $\tilde{G} = G(\tilde{\gamma}, \tilde{\delta})$, $\tilde{\theta} = \hat{\theta}(\tilde{\gamma}, \tilde{\delta})$, $\tilde{H} = H(\tilde{\delta})$, $\tilde{\phi} = \hat{\phi}(\tilde{\delta})$.
We have

$$\widehat{A} = \begin{bmatrix} \hat{a}_1' & \hat{a}_2 & \hat{a}_3 \\ 0' & \hat{a}_4 & \hat{a}_5 \\ 0' & 0 & \hat{a}_6 \end{bmatrix}, \tag{B.1}$$

where

$$\hat{a}_1' = 1'\tilde{G}^{-1}, \quad \hat{a}_2 = -1'\tilde{\theta}_c \tilde{s}_{cc}^{-1}, \tag{B.2}$$

$$\hat{a}_3 = 1'\tilde{\theta}_c \tilde{s}_{cc}^{-1} \tilde{s}_{cd} \tilde{s}_{dd}^{-1} - 1'\tilde{\theta}_d \tilde{s}_{dd}^{-1}, \quad \hat{a}_4 = -\tilde{s}_{cc}^{-1}, \tag{B.3}$$

$$\hat{a}_5 = \tilde{s}_{cc}^{-1} \tilde{s}_{cd} \tilde{s}_{dd}^{-1}, \quad \hat{a}_6 = -\tilde{s}_{dd}^{-1}, \tag{B.4}$$

in which

$$\tilde{\theta}_c = \tilde{G}^{-1}\left(\tilde{g}_c - \tilde{G}_c\tilde{\theta}\right), \quad \tilde{\theta}_d = \tilde{G}^{-1}\left(\tilde{g}_d - \tilde{G}_d\tilde{\theta}\right), \tag{B.5}$$

$$\tilde{g}_c = Q\frac{1}{n}\sum{}'\left\{Z_{tc}(\tilde{\gamma})y_t(\tilde{\gamma}) + Z_t(\tilde{\gamma},\tilde{\delta})y_{tc}(\tilde{\gamma})\right\}, \tag{B.6}$$

$$\tilde{G}_c = Q\frac{1}{n}\sum{}'\left\{Z_{tc}(\tilde{\gamma})Z_t'(\tilde{\gamma},\tilde{\delta}) + Z_t(\tilde{\gamma},\tilde{\delta})Z_{tc}'(\tilde{\gamma})\right\}Q', \tag{B.7}$$

$$\tilde{g}_d = Q\frac{1}{n}\sum{}'Z_{td}(\tilde{\delta})y_t(\tilde{\gamma}), \tag{B.8}$$

$$\tilde{G}_d = Q\frac{1}{n}\sum{}'\left\{Z_{td}(\tilde{\delta})Z_t'(\tilde{\gamma},\tilde{\delta}) + Z_t(\tilde{\gamma},\tilde{\delta})Z_{td}'(\tilde{\delta})\right\}Q', \tag{B.9}$$

with

$$y_{tc}(\tilde{\gamma}) = \log(1-L)y_t(\tilde{\gamma}), \tag{B.10}$$

$$Z_{tc}(\tilde{\gamma}) = \log(1-L)\left\{x_t(\tilde{\gamma}), 0, x_{t-1}(\tilde{\gamma}), 0, y_{t-1}(\tilde{\gamma}), ..., x_{t-p}(\tilde{\gamma}), 0, y_{t-p}(\tilde{\gamma})\right\}', \tag{B.11}$$

$$Z_{td}(\widetilde{\delta}) = \log(1-L)\left\{0, x_t(\widetilde{\delta}), 0, x_{t-1}(\widetilde{\delta}), 0, ..., 0, x_{t-p}(\widetilde{\delta}), 0\right\}', \tag{B.12}$$

and where

$$\widetilde{s}_{cc} = \frac{1}{n}\sum{}'\tilde{v}_{tc}^2, \quad \widetilde{s}_{cd} = \frac{1}{n}\sum{}'\tilde{v}_{tc}\tilde{v}_{td}, \quad \widetilde{s}_{dd} = \frac{1}{n}\sum{}'\tilde{w}_{td}^2, \tag{B.13}$$

with

$$\tilde{v}_{tc} = y_{tc}(\tilde{\gamma}) - \tilde{\theta}_c'QZ_t(\tilde{\gamma},\tilde{\delta}) - \tilde{\theta}'QZ_{tc}(\tilde{\gamma}), \tag{B.14}$$

$$\tilde{v}_{td} = -\tilde{\theta}_d'QZ_t(\tilde{\gamma},\tilde{\delta}) - \tilde{\theta}'QZ_{td}(\tilde{\delta}), \tag{B.15}$$

$$\tilde{w}_{td} = x_{td}(\tilde{\delta}) - \tilde{\phi}_d'RX_t(\tilde{\delta}) - \tilde{\phi}'RX_{td}(\tilde{\delta}), \tag{B.16}$$

$$x_{td}(\tilde{\delta}) = \log(1-L)x_t(\tilde{\delta}), \tag{B.17}$$

$$X_{td}(\tilde{\delta}) = \log(1-L)X_t(\tilde{d}), \tag{B.18}$$

$$\tilde{\phi}_d = \tilde{H}^{-1}(\tilde{h}_d - \tilde{H}_d\tilde{\phi}), \tag{B.19}$$

$$\tilde{h}_d = R\frac{1}{n}\sum{}'\left\{X_{td}(\tilde{\delta})x_t(\tilde{\delta}) + X_t(\tilde{\delta})x_{td}(\tilde{\delta})\right\}, \tag{B.20}$$

$$\tilde{H}_d = R\frac{1}{n}\sum{}'\left\{X_{td}(\tilde{\delta})X_t'(\tilde{\delta}) + X_t(\tilde{\delta})X_{td}'(\tilde{\delta})\right\}R'. \tag{B.21}$$

We also have

$$\widehat{B} = \frac{1}{n}\sum{}'\begin{bmatrix}\hat{\varepsilon}_{1.2,t}(\tilde{\gamma},\tilde{\delta})QZ_t(\tilde{\gamma},\tilde{\delta}) \\ \hat{\varepsilon}_{1.2,t}(\tilde{\gamma},\tilde{\delta})\tilde{v}_{tc} \\ \hat{\varepsilon}_{2t}(\tilde{\delta})\tilde{w}_{td}\end{bmatrix}\begin{bmatrix}\hat{\varepsilon}_{1.2,t}(\tilde{\gamma},\tilde{\delta})QZ_t(\tilde{\gamma},\tilde{\delta}) \\ \hat{\varepsilon}_{1.2,t}(\tilde{\gamma},\tilde{\delta})\tilde{v}_{tc} \\ \hat{\varepsilon}_{2t}(\tilde{\delta})\tilde{w}_{td}\end{bmatrix}', \tag{B.22}$$

where

$$\hat{\varepsilon}_{2t}(d) = x_t(d) - \tilde{\phi}'RX_t(d). \tag{B.23}$$

References

Box, G.E.P., Jenkins, G.M., 1971. Time series analysis. Forecasting and control. Holden-Day, San Francisco.

Brown, B.M.,1971. Martingale central limit theorems. Annals of Mathematical Statistics 42, 59-66.

Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models. Journal of Political Economy 95, 1062-1088.

Dolado J., Marmol, F., 1997. Efficient estimation of cointegrating relationships among higher order and fractionally integrated processes. Banco de Espana-Servicio de Estudios, Documento de Trabajo 9617.

Engle, R. F., Granger, C.W.J., 1987. Co-integration 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 processes. 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.

Hosoya, Y., 1997. Limit theory with long-range dependence and statistical inference of related models. Annals of Statistics 25, 105-137.

Jeganathan, P., 1999. On asymptotic inference in cointegrated time series with fractionally integrated errors. Econometric Theory 15, 583-621.

Jeganathan, P., 2001. Correction to 'On asymptotic inference in cointegrated time series with fractionally integrated errors'. Preprint.

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

Kashyap, R., Eom, K., 1988. Estimation in long memory time series model. Journal of Time Series Analysis 9, 35-41.

Kim, C.S., Phillips, P.C.B., 2000. Fully modified estimation of fractional cointegration models. Preprint.

Marinucci, D., Robinson, P.M., 2001. Semiparametric fractional cointegration analysis. Journal of Econometrics 105, 225-247.

Moulines, E., Soulier, P., 1999. Broadband log periodogram regression of time series with long range dependence. Annals of Statistics 27, 1415-1439.

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

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

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., Hualde, J., 2000. Cointegration in fractional systems with unknown integration orders. Preprint.

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. Annals of Statistics, forthcoming.

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.

Velasco, C., Robinson, P. M., 2000. Whittle pseudo-maximum likelihood estimation for nonstationary time series. Journal of the American Statistical Association 95, 1229-1243.

TABLE 2

MONTE CARLO BIAS, $b_1 = b_2 = 0$, correct specification

		n	64			128			256		
ρ	γ	δ	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$
	0	.4	-.006	-.005	-.007	-.001	-.001	-.003	-.001	-.002	.000
	.2	.4	-.014	-.036	-.011	.000	-.004	-.005	-.003	-.009	.000
0	.4	.8	-.006	-.002	-.015	-.001	-.002	-.009	-.001	-.001	-.002
	.7	1	-.009	-.024	-.031	.000	-.002	-.023	-.002	-.003	-.005
	0	.4	.001	-.117	.337	.005	-.032	.320	.003	-.009	.308
	.2	.4	-.001	-.268	.394	.009	-.143	.384	.006	-.071	.376
.5	.4	.8	.001	-.124	.192	.005	-.029	.155	.003	-.009	.120
	.7	1	.000	-.246	.214	.006	-.074	.182	.004	-.024	.143
	0	.4	.000	.104	-.338	-.002	.031	-.320	-.003	.007	-.307
	.2	.4	.000	.212	-.401	-.005	.137	-.387	-.010	.061	-.377
-.5	.4	.8	.000	.091	-.193	-.002	.027	-.151	-.003	.007	-.120
	.7	1	.000	.181	-.220	-.003	.065	-.176	-.006	.019	-.142
	0	.4	.002	-.178	.511	.003	-.042	.481	.002	-.011	.460
	.2	.4	.003	-.353	.599	.007	-.209	.578	.006	-.097	.562
.75	.4	.8	.002	-.177	.287	.003	-.043	.226	.002	-.010	.176
	.7	1	.003	-.308	.315	.005	-.120	.258	.004	-.031	.206

TABLE 3

MONTE CARLO BIAS, $b_1 = b_2 = 0.9$, correct specification

		n	64			128			256		
ρ	γ	δ	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$
	0	.4	-.026	-.150	-.014	-.016	.110	-.005	-.008	.025	.000
	.2	.4	-.057	.028	-.027	-.033	-.038	-.012	-.009	.013	-.001
0	.4	.8	-.026	.019	-.025	-.016	.052	-.014	-.008	-.011	-.003
	.7	1	-.036	-.012	-.043	-.022	-.153	-.030	-.008	.001	-.006
	0	.4	.016	.050	.158	.004	-.023	.137	.005	-.003	.120
	.2	.4	.028	-.094	.281	.010	.135	.267	.008	.086	.247
.5	.4	.8	.016	-.109	.140	.004	-.052	.116	.005	-.020	.090
	.7	1	.019	-.287	.195	.006	-.191	.170	.006	-.034	.134
	0	.4	-.015	-.001	-.161	-.003	-.025	-.136	-.005	.010	-.120
	.2	.4	-.041	.130	-.293	-.008	-.023	-.266	-.006	-.140	-.248
-.5	.4	.8	-.015	.065	-.147	-.003	.024	-.113	-.005	.040	-.088
	.7	1	-.024	.299	-.207	-.005	.121	-.166	-.006	.136	-.131
	0	.4	.027	.037	.237	.010	-.025	.202	.007	.018	.176
	.2	.4	.047	-.025	.421	.020	.093	.390	.010	.134	.364
.75	.4	.8	.027	-.194	.206	.010	-.038	.165	.007	.005	.129
	.7	1	.034	-.483	.283	.013	-.270	.236	.008	-.116	.192

TABLE 4

MONTE CARLO BIAS, $b_1 = 0$, $b_2 = 0.5$, correct specification

		n	64			128			256		
ρ	γ	δ	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$
	0	.4	-.001	-.003	-.004	.001	.004	-.001	.001	.001	.000
	.2	.4	.001	-.016	-.008	.004	-.001	-.003	.003	.009	.000
0	.4	.8	-.001	-.022	-.008	.001	.005	-.005	.001	.001	-.001
	.7	1	.000	-.044	-.017	.002	.012	-.012	.002	-.001	-.002
	0	.4	.006	.009	.142	.004	-.003	.129	.001	.001	.119
	.2	.4	.016	.028	.201	.010	-.013	.189	.004	.000	.180
.5	.4	.8	.006	.010	.082	.004	.001	.067	.001	.001	.052
	.7	1	.009	.002	.102	.006	-.006	.088	.002	-.004	.069
	0	.4	-.001	.001	-.142	.000	.005	-.128	.000	.004	-.119
	.2	.4	-.002	-.031	-.203	.001	.011	-.189	-.001	.021	-.181
-.5	.4	.8	-.001	-.003	-.083	.000	.008	-.065	.000	.004	-.052
	.7	1	-.001	-.009	-.106	.000	.015	-.085	.000	.017	-.069
	0	.4	.004	.005	.216	.002	.002	.192	.000	.000	.178
	.2	.4	.011	.042	.305	.006	-.004	.283	.001	-.017	.269
.75	.4	.8	.004	.002	.123	.002	.000	.097	.000	.001	.076
	.7	1	.006	-.012	.151	.003	-.018	.124	.001	-.010	.100

TABLE 5

MONTE CARLO BIAS, $b_1 = 0.9$, $b_2 = 0$, correct specification

		n	64			128			256		
ρ	γ	δ	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$
	0	.4	.006	-.056	-.039	.005	.045	-.015	.005	.013	-.002
	.2	.4	-.002	-.056	-.065	.009	.020	-.030	.005	-.001	-.005
0	.4	.8	.006	-.053	-.119	.005	.053	-.082	.005	.014	-.013
	.7	1	.003	-.063	-.251	.006	.039	-.210	.005	.009	-.038
	0	.4	.129	.325	.714	.052	.165	.740	.018	.050	.741
	.2	.4	.258	.223	.970	.126	.141	1.07	.056	.061	1.12
.5	.4	.8	.129	.333	.994	.052	.167	.981	.018	.055	.854
	.7	1	.177	.240	1.42	.079	.148	1.46	.032	.053	1.27
	0	.4	-.118	-.457	-.758	-.040	-.144	-.755	-.014	-.043	-.746
	.2	.4	-.264	-.403	-1.05	-.110	-.153	-1.11	-.054	-.094	-1.14
-.5	.4	.8	-.118	-.475	-1.05	-.040	-.143	-.965	-.014	-.045	-.852
	.7	1	-.172	-.397	-1.51	-.066	-.159	-1.41	-.029	-.068	-1.26
	0	.4	.167	.419	1.09	.065	.192	1.11	.022	.036	1.11
	.2	.4	.363	.379	1.48	.172	.213	1.61	.079	.064	1.68
.75	.4	.8	.167	.423	1.48	.065	.191	1.42	.022	.036	1.25
	.7	1	.242	.376	2.08	.106	.166	2.05	.043	.049	1.83

TABLE 6

MONTE CARLO BIAS, $b_1 = b_2 = 0.5$, mis-specification

ρ	γ	n / δ	64 $\overline{\nu}_I$	64 $\overline{\nu}_F$	64 $\overline{\nu}_0$	128 $\overline{\nu}_I$	128 $\overline{\nu}_F$	128 $\overline{\nu}_0$	256 $\overline{\nu}_I$	256 $\overline{\nu}_F$	256 $\overline{\nu}_0$
0	0	.4	-.005	.003	-.008	.000	.000	-.003	.000	-.004	.000
	.2	.4	-.010	-.027	-.016	.002	-.002	-.006	-.001	-.015	-.001
	.4	.8	-.005	.009	-.017	.000	.001	-.010	.000	-.002	-.002
	.7	1	-.007	-.004	-.033	.000	.005	-.024	.000	-.015	-.005
.5	0	.4	.004	-.240	.240	.006	-.188	.222	.003	-.096	.208
	.2	.4	.008	-.361	.337	.013	-.365	.326	.007	-.343	.314
	.4	.8	.004	-.352	.164	.006	-.230	135	.003	-.140	.105
	.7	1	.005	-.808	.204	.008	-.842	177	.004	-.866	.140
-.5	0	.4	.000	.176	-.242	-.001	.174	-.221	-.003	.101	-.208
	.2	.4	.000	.287	-.346	-.003	.356	-.328	-.009	.304	-.316
	.4	.8	.000	.299	-.167	-.001	.244	-.132	-.003	.146	-.105
	.7	1	.000	.790	-.212	-.002	.818	-.170	-.005	.883	-.138
.75	0	.4	.004	-.318	.365	.003	-.217	.332	.002	-.117	.310
	.2	.4	.009	-.500	.513	.008	-.564	.487	.006	-.493	.469
	.4	.8	.004	-.457	.244	.003	-.280	.196	.002	-.154	.154
	.7	1	.006	-1.20	.300	.005	-1.18	.250	.003	-1.18	.201

TABLE 7

MONTE CARLO BIAS, $b_1 = b_2 = 0$, over-specification

ρ	γ	n / δ	64 $\overline{\nu}_I$	64 $\overline{\nu}_F$	64 $\overline{\nu}_0$	128 $\overline{\nu}_I$	128 $\overline{\nu}_F$	128 $\overline{\nu}_0$	256 $\overline{\nu}_I$	256 $\overline{\nu}_F$	256 $\overline{\nu}_0$
0	0	.4	-.032	-.138	-.007	-.006	.013	-.003	.007	.006	.000
	.2	.4	-.036	.007	-.011	.023	.021	-.005	.027	.024	.000
	.4	.8	-.032	-.091	-.015	-.006	.000	-.009	.007	-.001	-.002
	.7	1	-.034	-.036	-.031	.003	.000	-.023	.014	-.009	-.005
.5	0	.4	.006	.040	.337	.017	.044	.320	-.005	.014	.308
	.2	.4	.021	-.122	.394	.061	.020	.384	.004	-.036	.376
	.4	.8	.006	.019	.192	.017	.021	.155	-.005	.007	.120
	.7	1	.012	-.133	.214	.032	.043	.182	-.001	.008	.143
-.5	0	.4	.020	-.047	-.338	.013	.032	-.320	.021	.018	-.307
	.2	.4	.065	-.129	-.401	.042	.137	-.387	.035	.086	-.377
	.4	.8	.020	-.053	-.193	.013	.045	-.151	.021	.033	-.120
	.7	1	.035	-.044	-.220	.022	.028	-.176	.026	.062	-.142
.75	0	.4	-.018	.002	.511	.002	.086	.481	-.016	.001	.460
	.2	.4	-.034	-.083	.599	.016	-.127	.578	-.021	-.124	.562
	.4	.8	-.018	-.016	.287	.002	.058	.226	-.016	-.013	.176
	.7	1	-.023	-.118	.315	.007	-.051	.258	-.017	-.037	.206

TABLE 8

MONTE CARLO S.D., $b_1 = b_2 = 0$, correct specification

ρ	γ	n / δ	64 $\bar{\nu}_I$	$\bar{\nu}_F$	$\bar{\nu}_0$	128 $\bar{\nu}_I$	$\bar{\nu}_F$	$\bar{\nu}_0$	256 $\bar{\nu}_I$	$\bar{\nu}_F$	$\bar{\nu}_0$
0	0	.4	.212	.383	.107	.128	.160	.073	.086	.092	.049
	.2	.4	.489	1.03	.141	.310	.559	.105	.217	.318	.076
	.4	.8	.212	.387	.171	.128	.159	.128	.086	.092	.093
	.7	1	.305	.679	.322	.189	.323	.278	.130	.153	.214
.5	0	.4	.184	.566	.112	.113	.218	.084	.073	.098	.063
	.2	.4	.426	1.12	.136	.276	.650	.104	.187	.366	.078
	.4	.8	.184	.569	.160	.113	.194	.127	.073	.098	.092
	.7	1	.266	.913	.283	.168	.376	.247	.112	.176	.192
-.5	0	.4	.178	.528	.109	.112	.227	.084	.076	.101	.065
	.2	.4	.419	1.01	.131	.274	.614	.102	.193	.359	.077
	.4	.8	.178	.485	.154	.112	.221	.122	.076	.103	.092
	.7	1	.259	.758	.270	.167	.361	.237	.116	.185	.188
.75	0	.4	.140	.711	.114	.087	.237	.091	.058	.102	.075
	.2	.4	.328	1.08	.116	.213	.706	.092	.146	.426	.073
	.4	.8	.140	.734	.140	.087	.260	.111	.058	.101	.086
	.7	1	.203	.973	.226	.129	.537	.188	.088	.197	.152

TABLE 9

MONTE CARLO S.D., $b_1 = b_2 = 0.9$, correct specification

ρ	γ	n / δ	64 $\bar{\nu}_I$	$\bar{\nu}_F$	$\bar{\nu}_0$	128 $\bar{\nu}_I$	$\bar{\nu}_F$	$\bar{\nu}_0$	256 $\bar{\nu}_I$	$\bar{\nu}_F$	$\bar{\nu}_0$
0	0	.4	1.06	4.16	.192	.553	2.10	.122	.306	1.43	.079
	.2	.4	2.04	5.28	.354	1.10	3.06	.253	.634	1.93	.177
	.4	.8	1.06	4.14	.282	.553	2.31	.191	.306	1.24	.120
	.7	1	1.37	4.22	.483	.729	2.10	.370	.411	1.01	.249
.5	0	.4	.901	3.38	.172	.472	2.18	.115	.266	1.23	.075
	.2	.4	1.76	4.46	.319	.953	3.01	.233	.553	1.65	.161
	.4	.8	.901	3.51	.241	.472	2.15	.170	.266	1.19	.109
	.7	1	1.17	4.10	.405	.625	2.39	.313	.358	1.02	.219
-.5	0	.4	.918	3.47	.164	.480	1.93	.112	.271	1.16	.075
	.2	.4	1.78	5.12	.300	.961	2.85	.225	.557	1.67	.159
	.4	.8	.918	3.73	.225	.480	1.90	.161	.271	1.09	.108
	.7	1	1.19	3.82	.374	.633	2.15	.296	.363	1.26	.216
.75	0	.4	.717	2.75	.138	.372	1.67	.093	.212	.946	.066
	.2	.4	1.39	3.93	.248	.747	2.26	.179	.441	1.37	.131
	.4	.8	.717	3.22	.195	.372	1.52	.128	.212	.823	.088
	.7	1	.930	3.38	.331	.491	1.83	.232	.286	.938	.169

TABLE 10

MONTE CARLO S.D., $b_1 = 0.9$, $b_2 = 0$, correct specification

ρ	γ	n / δ	64 $\overline{\nu}_I$	64 $\overline{\nu}_F$	64 $\overline{\nu}_0$	128 $\overline{\nu}_I$	128 $\overline{\nu}_F$	128 $\overline{\nu}_0$	256 $\overline{\nu}_I$	256 $\overline{\nu}_F$	256 $\overline{\nu}_0$
0	0	.4	.666	2.08	.466	.399	1.22	.373	.239	.560	.280
	.2	.4	1.14	2.65	.864	.711	1.66	.764	.443	.882	.615
	.4	.8	.666	2.05	1.36	.399	1.20	1.18	.239	.567	.907
	.7	1	.813	2.24	2.70	.496	1.33	2.60	.301	.655	2.09
.5	0	.4	.615	2.06	.450	.358	1.09	.353	.205	.468	.262
	.2	.4	1.09	2.55	.849	.657	1.62	.729	.408	.816	.585
	.4	.8	.615	2.15	1.24	.358	1.10	1.08	.205	.486	.816
	.7	1	.768	2.25	2.38	.451	1.28	2.27	.268	.586	1.85
-.5	0	.4	.608	2.16	.434	.346	1.13	.338	.210	.510	.249
	.2	.4	1.09	2.72	.831	.642	1.51	.714	.403	.903	.555
	.4	.8	.608	2.21	1.18	.346	1.14	1.04	.210	.515	.818
	.7	1	.761	2.34	2.23	.438	1.34	2.16	.270	.678	1.81
.75	0	.4	.529	2.01	.383	.295	1.02	.297	.166	.349	.217
	.2	.4	.986	2.67	.769	.590	1.55	.652	.362	.787	.508
	.4	.8	.529	2.05	.974	.295	.941	.835	.166	.359	.678
	.7	1	.681	2.24	1.89	.391	1.13	1.72	.228	.521	1.44

TABLE 11

MONTE CARLO S.D., $b_1 = b_2 = 0$, over-specification

ρ	γ	n / δ	64 $\overline{\nu}_I$	64 $\overline{\nu}_F$	64 $\overline{\nu}_0$	128 $\overline{\nu}_I$	128 $\overline{\nu}_F$	128 $\overline{\nu}_0$	256 $\overline{\nu}_I$	256 $\overline{\nu}_F$	256 $\overline{\nu}_0$
0	0	.4	2.04	4.46	.107	1.19	2.23	.073	.748	.929	.049
	.2	.4	4.03	6.98	.141	2.37	4.40	.105	1.52	2.70	.076
	.4	.8	2.04	4.40	.171	1.19	2.24	.128	.748	.914	.093
	.7	1	2.66	5.39	.322	1.56	3.38	.278	.988	1.59	.214
.5	0	.4	1.74	3.22	.112	1.06	1.79	.084	.668	.907	.063
	.2	.4	3.39	6.04	.136	2.12	3.90	.104	1.35	2.39	.078
	.4	.8	1.74	3.47	.160	1.06	1.72	.127	.668	.899	.092
	.7	1	2.26	4.53	.283	1.40	2.85	.247	.881	1.44	.192
-.5	0	.4	1.78	3.55	.109	1.07	1.91	.084	.670	.925	.065
	.2	.4	3.46	5.48	.131	2.14	3.92	.102	1.36	2.33	.077
	.4	.8	1.78	3.42	.154	1.07	1.92	.122	.670	.971	.092
	.7	1	2.30	4.52	.270	1.41	2.85	.237	.887	1.53	.188
.75	0	.4	1.42	2.73	.114	.831	1.63	.091	.519	.651	.075
	.2	.4	2.74	4.51	.116	1.67	3.24	.092	1.05	1.85	.073
	.4	.8	1.42	2.76	.140	.831	1.57	.111	.519	.636	.086
	.7	1	1.83	3.48	.226	1.09	2.31	.188	.686	1.06	.152

TABLE 12

EMPIRICAL SIZES OF W_I AND W_F, $b_1 = b_2 = 0$, correct specification

		α	.05						.10					
		n	64	64	128	128	256	256	64	64	128	128	256	256
ρ	γ	δ	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F
	0	.4	.078	.061	.053	.056	.057	.059	.136	.122	.112	.094	.125	.114
	.2	.4	.077	.045	.054	.032	.062	.034	.133	.083	.104	.072	.114	.069
0	.4	.8	.078	.059	.053	.055	.057	.059	.136	.125	.112	.087	.125	.114
	.7	1	.076	.058	.058	.057	.053	.053	.134	.107	.105	.103	.120	.098
	0	.4	.074	.057	.055	.061	.055	.065	.136	.089	.119	.092	.117	.111
	.2	.4	.073	.105	.055	.082	.054	.079	.141	.153	.120	.128	.111	.112
.5	.4	.8	.074	.059	.055	.057	.055	.066	.136	.089	.119	.094	.117	.111
	.7	1	.068	.088	.055	.076	.050	.069	.140	.125	.121	.117	.116	.109
	0	.4	.076	.063	.072	.061	.068	.068	.124	.103	.124	.107	.122	.118
	.2	.4	.076	.123	.059	.106	.058	.084	.134	.168	.117	.145	.130	.119
-.5	.4	.8	.076	.071	.072	.059	.068	.069	.124	.101	.124	.105	.122	.118
	.7	1	.073	.102	.066	.086	.060	.078	.129	.144	.118	.142	.128	.117
	0	.4	.075	.052	.059	.054	.063	.070	.136	.083	.112	.097	.116	.111
	.2	.4	.073	.168	.058	.136	.069	.094	.143	.207	.113	.166	.116	.132
.75	.4	.8	.075	.049	.059	.054	.063	.073	.136	.083	.112	.097	.116	.110
	.7	1	.076	.120	.060	.105	.064	.078	.143	.155	.113	.138	.110	.117

TABLE 13

EMPIRICAL SIZES OF W_I AND W_F, $b_1 = b_2 = 0.9$, correct specification

		α	.05						.10					
		n	64	64	128	128	256	256	64	64	128	128	256	256
ρ	γ	δ	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F
	0	.4	.122	.038	.080	.035	.077	.025	.187	.066	.150	.064	.129	.053
	.2	.4	.125	.033	.092	.023	.063	.024	.191	.069	.146	.051	.130	.050
0	.4	.8	.122	.032	.080	.033	.077	.033	.187	.068	.150	.068	.129	.064
	.7	1	.125	.043	.079	.035	.075	.018	.192	.073	.146	.055	.122	.055
	0	.4	.112	.027	.097	.031	.067	.030	.177	.054	.160	.064	.145	.063
	.2	.4	.118	.035	.094	.042	.071	.055	.182	.069	.161	.084	.139	.096
.5	.4	.8	.112	.038	.097	.035	.067	.036	.177	.080	.160	.069	.145	.064
	.7	1	.121	.048	.090	.039	.073	.055	.179	.075	.165	.070	.133	.081
	0	.4	.114	.037	.092	.034	.084	.028	.184	.080	.161	.071	.132	.063
	.2	.4	.109	.048	.098	.046	.074	.054	.180	.088	.158	.088	.138	.101
-.5	.4	.8	.114	.054	.092	.039	.084	.036	.184	.079	.161	.070	.132	.068
	.7	1	.112	.060	.097	.044	.082	.053	.182	.089	.161	.072	.136	.093
	0	.4	.115	.035	.100	.026	.079	.035	.185	.069	.161	.069	.151	.059
	.2	.4	.107	.057	.096	.063	.081	.105	.188	.108	.162	.104	.146	.156
.75	.4	.8	.115	.047	.100	.033	.079	.033	.185	.073	.161	.062	.151	.061
	.7	1	.112	.046	.101	.059	.079	.061	.181	.090	.159	.087	.141	.106

TABLE 14

EMPIRICAL SIZES OF W_I AND W_F, $b_1 = 0$, $b_2 = 0.5$, correct specification

		α	.05						.10					
		n	64	64	128	128	256	256	64	64	128	128	256	256
ρ	γ	δ	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F
	0	.4	.069	.010	.067	.022	.059	.028	.113	.018	.122	.048	.106	.072
	.2	.4	.066	.020	.064	.023	.065	.018	.114	.035	.120	.041	.112	.038
0	.4	.8	.069	.010	.067	.017	.059	.029	.113	.018	.122	.050	.106	.068
	.7	1	.070	.015	.067	.027	.065	.023	.114	.034	.125	.054	.107	.062
	0	.4	.062	.020	.054	.024	.049	.034	.124	.042	.115	.053	.105	.054
	.2	.4	.061	.044	.053	.064	.049	.059	.127	.078	.110	.091	.103	.096
.5	.4	.8	.062	.019	.054	.022	.049	.037	.124	.039	.115	.051	.105	.057
	.7	1	.066	.040	.051	.045	.047	.054	.127	.076	.118	.069	.102	.076
	0	.4	.067	.017	.067	.018	.055	.033	.125	.033	.117	.045	.100	.059
	.2	.4	.067	.053	.063	.063	.055	.059	.119	.082	.119	.095	.094	.088
-.5	.4	.8	.067	.013	.067	.019	.055	.031	.125	.035	.117	.046	.100	.054
	.7	1	.067	.045	.066	.038	.058	.047	.122	.071	.120	.074	.103	.073
	0	.4	.073	.024	.055	.025	.054	.022	.145	.037	.107	.053	.096	.043
	.2	.4	.069	.108	.054	.126	.057	.113	.131	.158	.104	.164	.099	.151
.75	.4	.8	.073	.031	.055	.024	.054	.023	.145	.051	.107	.056	.096	.051
	.7	1	.067	.082	.058	.055	.051	.065	.137	.117	.106	.096	.103	.106

TABLE 15

EMPIRICAL SIZES OF W_I AND W_F, $b_1 = 0.9$, $b_2 = 0$, correct specification

		α	.05						.10					
		n	64	64	128	128	256	256	64	64	128	128	256	256
ρ	γ	δ	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F
	0	.4	.097	.053	.086	.042	.071	.042	.162	.087	.157	.077	.125	.072
	.2	.4	.090	.038	.091	.026	.077	.021	.166	.065	.150	.045	.127	.042
0	.4	.8	.097	.044	.086	.042	.071	.045	.162	.080	.157	.075	.125	.073
	.7	1	.092	.039	.089	.035	.070	.030	.155	.068	.150	.068	.124	.056
	0	.4	.112	.041	.073	.031	.053	.031	.165	.063	.141	.059	.101	.066
	.2	.4	.097	.023	.078	.027	.064	.019	.161	.045	.139	.049	.120	.045
.5	.4	.8	.112	.043	.073	.030	.053	.031	.165	.063	.141	.058	.101	.062
	.7	1	.109	.027	.082	.031	.060	.032	.164	.054	.147	.064	.110	.064
	0	.4	.101	.051	.081	.033	.068	.030	.171	.082	.140	.062	.115	.062
	.2	.4	.105	.031	.087	.023	.060	.021	.178	.059	.139	.046	.123	.031
-.5	.4	.8	.101	.051	.081	.034	.068	.030	.171	.081	.140	.060	.115	.061
	.7	1	.101	.036	.086	.031	.061	.031	.175	.068	.140	.052	.119	.055
	0	.4	.117	.032	.082	.024	.051	.021	.185	.053	.133	.052	.104	.051
	.2	.4	.107	.028	.078	.026	.065	.024	.173	.044	.133	.042	.114	.043
.75	.4	.8	.117	.033	.082	.022	.051	.021	.185	.053	.133	.053	.104	.051
	.7	1	.111	.030	.081	.028	.058	.029	.184	.059	.143	.053	.106	.054

TABLE 16

EMPIRICAL SIZES OF W_I AND W_F, $b_1 = b_2 = 0.5$, mis-specification

		α			.05						.10			
		n	64	64	128	128	256	256	64	64	128	128	256	256
ρ	γ	δ	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F
	0	.4	.258	.026	.245	.027	.248	.037	.344	.063	.319	.060	.325	.079
	.2	.4	.242	.013	.214	.012	.229	.013	.327	.043	.296	.024	.310	.035
0	.4	.8	.258	.022	.245	.024	.248	.035	.344	.061	.319	.052	.325	.073
	.7	1	.255	.019	.229	.006	.241	.017	.339	.042	.308	.029	.322	.031
	0	.4	.264	.040	.246	.030	.248	.032	.356	.070	.324	.052	.324	.069
	.2	.4	.245	.072	.230	.051	.224	.052	.341	.105	.303	.079	.317	.064
.5	.4	.8	.264	.033	.246	.028	.248	.029	.356	.054	.324	.046	.324	.070
	.7	1	.253	.031	.239	.028	.239	.014	.347	.047	.306	.043	.325	.025
	0	.4	.274	.033	.250	.026	.255	.030	.349	.067	.333	.053	.341	.068
	.2	.4	.258	.077	.228	.053	.228	.047	.331	.117	.317	.080	.317	.073
-.5	.4	.8	.274	.031	.250	.024	.255	.024	.349	.058	.333	.046	.341	.070
	.7	1	.270	.036	.243	.019	.233	.011	.343	.050	.331	.033	.334	.022
	0	.4	.274	.035	.244	.024	.251	.025	.360	.057	.329	.043	.333	.064
	.2	.4	.249	.119	.221	.079	.218	.054	.336	.155	.310	.099	.313	.071
.75	.4	.8	.274	.028	.244	.022	.251	.025	.360	.044	.329	.034	.333	.063
	.7	1	.262	.041	.240	.032	.238	.010	.350	.051	.318	.040	.318	.013

TABLE 17

EMPIRICAL SIZES OF W_I AND W_F, $b_1 = b_2 = 0$, over-specification

		α			.05						.10			
		n	64	64	128	128	256	256	64	64	128	128	256	256
ρ	γ	δ	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F	W_I	W_F
	0	.4	.078	.047	.061	.047	.050	.042	.127	.085	.115	.088	.100	.082
	.2	.4	.072	.040	.054	.042	.047	.027	.135	.075	.107	.070	.086	.048
0	.4	.8	.078	.049	.061	.041	.050	.042	.127	.091	.115	.083	.100	.080
	.7	1	.075	.049	.052	.037	.049	.033	.132	.083	.104	.074	.094	.074
	0	.4	.068	.037	.063	.052	.056	.048	.124	.071	.118	.093	.105	.082
	.2	.4	.071	.052	.064	.045	.061	.026	.113	.079	.116	.071	.110	.046
.5	.4	.8	.068	.039	.063	.050	.056	.047	.124	.071	.118	.088	.105	.083
	.7	1	.065	.043	.056	.047	.060	.046	.120	.076	.110	.087	.110	.074
	0	.4	.091	.057	.072	.048	.066	.048	.143	.087	.109	.093	.112	.095
	.2	.4	.084	.051	.065	.049	.053	.021	.139	.088	.115	.080	.099	.056
-.5	.4	.8	.091	.054	.072	.051	.066	.051	.143	.092	.109	.090	.112	.100
	.7	1	.088	.062	.067	.051	.058	.040	.137	.103	.112	.094	.105	.084
	0	.4	.085	.052	.072	.047	.060	.047	.144	.087	.129	.081	.113	.085
	.2	.4	.074	.051	.073	.057	.057	.026	.138	.099	.126	.084	.114	.045
.75	.4	.8	.085	.049	.072	.042	.060	.047	.144	.084	.129	.076	.113	.088
	.7	1	.080	.056	.080	.051	.058	.044	.143	.093	.125	.093	.112	.090

TABLE 18

MONTE CARLO BIAS of $\widetilde{\delta}$, $\rho = 0.5$

estimation	n / $\delta \backslash b_2$	64: 0	64: .5	64: .9	128: 0	128: .5	128: .9	256: 0	256: .5	256: .9
$S_0(d)$	.4	-.023	.377	.795	-.011	.358	.818	-.005	.363	.833
	.8	-.025	.328	.493	-.008	.332	.524	-.004	.343	.545
	1	-.036	.227	.267	-.014	.232	.292	-.006	.236	.290
$S_1(d)$	.4	-.045	.127	.662	-.029	.047	.595	-.015	.025	.570
	.8	-.040	.105	.405	-.017	.048	.379	-.011	.029	.356
	1	-.051	.047	.196	-.033	.015	.166	-.016	.007	.150

TABLE 19

MONTE CARLO S.D. of $\widetilde{\delta}$, $\rho = 0.5$

estimation	n / $\delta \backslash b_2$	64: 0	64: .5	64: .9	128: 0	128: .5	128: .9	256: 0	256: .5	256: .9
$S_0(d)$	.4	.125	.135	.139	.082	.105	.104	.052	.073	.071
	.8	.125	.145	.206	.082	.110	.203	.051	.079	.193
	1	.122	.164	.217	.079	.139	.222	.050	.113	.211
$S_1(d)$	.4	.253	.240	.259	.161	.171	.222	.093	.116	.172
	.8	.257	.245	.275	.170	.174	.254	.095	.119	.232
	1	.240	.224	.278	.163	.168	.249	.092	.116	.221

TABLE 20

EMPIRICAL SIZES ($\alpha = 0.05$) OF W_δ, $\rho = 0.5$

estimation	n / $\delta \backslash b_2$	64: 0	64: .5	64: .9	128: 0	128: .5	128: .9	256: 0	256: .5	256: .9
$S_0(d)$	.4	.134	.902	1.00	.099	.968	1.00	.073	1.00	1.00
	.8	.126	.839	.984	.095	.952	.993	.068	.997	1.00
	1	.121	.611	.786	.082	.800	.923	.064	.918	.981
$S_1(d)$	.4	.129	.140	.685	.103	.084	.741	.074	.063	.877
	.8	.123	.125	.337	.115	.090	.424	.080	.058	.473
	1	.088	.083	.141	.090	.048	.146	.069	.035	.177

TABLE 21

EMPIRICAL SIZES ($\alpha = 0.10$) OF W_δ, $\rho = 0.5$

estimation	n / $\delta \backslash b_2$	64: 0	64: .5	64: .9	128: 0	128: .5	128: .9	256: 0	256: .5	256: .9
$S_0(d)$	.4	.188	.935	1.00	.147	.975	1.00	.122	1.00	1.00
	.8	.191	.889	.989	.151	.970	.996	.123	.997	1.00
	1	.177	.705	.851	.136	.856	.939	.111	.941	.983
$S_1(d)$	.4	.190	.190	.752	.175	.127	.792	.129	.091	.930
	.8	.186	.168	.397	.187	.129	.479	.137	.099	.529
	1	.150	.116	.158	.150	.088	.182	.130	.066	.210

TABLE 22

MONTE CARLO BIAS of $\widetilde{\gamma}$, $\rho = 0.5$, $b_1 = b_2 = b$

		n	64			128			256		
estimation	γ	$\delta\backslash^{b}$	0	.5	.9	0	.5	.9	0	.5	.9
$S_0\,(d)$, $T_0\,(c)$	0	.4	-.008	.420	.857	-.006	.413	.866	-.004	.414	.871
	.2	.4	-.046	.372	.809	-.020	.388	.844	-.006	.407	.865
	.4	.8	-.008	.405	.743	-.005	.409	.775	-.004	.413	.794
	.7	1	-.034	.347	.494	-.015	.371	.520	-.006	.389	.523
$S_1\,(d)$, $T_1\,(c)$	0	.4	-.047	.094	.642	-.027	.024	.582	-.015	.001	.561
	.2	.4	-.176	-.051	.481	-.103	-.098	.412	-.040	-.089	.376
	.4	.8	-.043	.079	.414	-.020	.026	.387	-.013	.003	.343
	.7	1	-.116	-.042	.173	-.070	-.058	.149	-.032	-.056	.115

TABLE 23

MONTE CARLO S.D. of $\widetilde{\gamma}$, $\rho = 0.5$, $b_1 = b_2 = b$

		n	64			128			256		
estimation	γ	$\delta\backslash^{b}$	0	.5	.9	0	.5	.9	0	.5	.9
$S_0\,(d)$, $T_0\,(c)$	0	.4	.096	.106	.107	.067	.078	.079	.048	.057	.056
	.2	.4	.106	.111	.111	.075	.085	.085	.051	.058	.058
	.4	.8	.095	.110	.124	.066	.081	.107	.048	.058	.090
	.7	1	.103	.113	.173	.074	.087	.177	.051	.061	.172
$S_1\,(d)$, $T_1\,(c)$	0	.4	.233	.220	.254	.133	.159	.224	.077	.115	.192
	.2	.4	.266	.224	.270	.180	.179	.253	.094	.151	.233
	.4	.8	.232	.220	.288	.142	.155	.265	.077	.114	.267
	.7	1	.237	.216	.277	.159	.162	.241	.089	.134	.216

TABLE 24

EMPIRICAL SIZES ($\alpha = 0.05$) OF W_γ, $\rho = 0.5$, $b_1 = b_2 = b$

		n	64			128			256		
estimation	γ	$\delta\backslash^{b}$	0	.5	.9	0	.5	.9	0	.5	.9
$S_0\,(d)$, $T_0\,(c)$	0	.4	.038	.979	1.00	.041	1.00	1.00	.051	1.00	1.00
	.2	.4	.087	.939	1.00	.085	.995	1.00	.059	1.00	1.00
	.4	.8	.039	.969	1.00	.040	1.00	1.00	.045	1.00	1.00
	.7	1	.064	.917	.993	.076	.992	.998	.059	1.00	1.00
$S_1\,(d)$, $T_1\,(c)$	0	.4	.072	.098	.683	.039	.058	.722	.034	.033	.826
	.2	.4	.096	.053	.455	.106	.077	.464	.090	.092	.505
	.4	.8	.070	.084	.383	.047	.055	.412	.034	.032	.465
	.7	1	.070	.061	.163	.060	.045	.152	.074	.058	.161

TABLE 25

EMPIRICAL SIZES ($\alpha = 0.10$) OF W_γ, $\rho = 0.5$, $b_1 = b_2 = b$

		n	64			128			256		
estimation	γ	$\delta\backslash^{b}$	0	.5	.9	0	.5	.9	0	.5	.9
$S_0\,(d)$, $T_0\,(c)$	0	.4	.075	.993	1.00	.081	1.00	1.00	.107	1.00	1.00
	.2	.4	.151	.964	1.00	.128	.997	1.00	.111	1.00	1.00
	.4	.8	.082	.984	1.00	.080	1.00	1.00	.108	1.00	1.00
	.7	1	.122	.946	.997	.123	.995	.998	.123	1.00	1.00
$S_1\,(d)$, $T_1\,(c)$	0	.4	.116	.135	.733	.080	.088	.778	.064	.068	.890
	.2	.4	.157	.094	.515	.179	.113	.520	.151	.143	.569
	.4	.8	.120	.127	.427	.094	.087	.461	.062	.068	.512
	.7	1	.112	.083	.191	.121	.075	.189	.129	.092	.195