ROOT-N-CONSISTENT ESTIMATION OF WEAK FRACTIONAL COINTEGRATION

J. Hualdea[*], P.M. Robinson[§]._

The Suntory Centre
Suntory and Toyota International Centres for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel: 020 7955 6679

Discussion paper
No. EM/06/499
March2006

[*] School of Economics and Business Administration, Universidad de Navarra, Pamplona 31080, Spain

[§] Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, UK

Abstract

Empirical evidence has emerged of the possibility of fractional cointegration such that the gap, β, between the integration order δ of observable time series, and the integration order γ of cointegrating errors, is less than 0.5. This includes circumstances when observables are stationary or asymptotically stationary with long memory (so $\delta < 1/2$), and when they are nonstationary (so $\delta \geq 1/2$). This "weak cointegration" contrasts strongly with the traditional econometric prescription of unit root observables and short memory cointegrating errors, where $\beta = 1$. Asymptotic inferential theory also differs from this case, and from other members of the class $\beta > 1/2$, in particular $\sqrt{n}$-consistent and asymptotically normal estimation of the cointegrating vector ν is possible when $\beta < 1/2$, 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.

JEL Classification: C32.
Keywords: Fractional cointegration; Parametric estimation; Asymptotic normality.

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

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

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 zero mean and 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 $CI(\delta, \gamma)$ (Engle and Granger, 1987), for which it is necessary that y_t and x_t share the same integration order (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, whereas, for example, $\Delta^{-\delta} u_{2t}$ does not have finite variance when $\delta \geq 1/2$.

We anticipate

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

when, rewriting the first equation of (1) as the regression

$$y_t = \nu x_t + v_{1t}, \;\; v_{1t} = \Delta^{-\gamma} u_{1t}^{\#}, \tag{7}$$

the x_t and v_{1t} are contemporaneously correlated. When

$$\delta < 1/2 \tag{8}$$

(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 v_{1t}. To overcome this problem, Robinson (1994a) showed that a narrow-band frequency-domain least squares estimate (NBLSE) is consistent, due to 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.) Robinson and Marinucci (2003) gave a rate of convergence for this latter estimate, conjecturing its sharpness. Assuming (4), (8) and $\gamma + \delta < 1/2$, Christensen and Nielsen (2004) obtained the asymptotic distribution of the NBLSE when u_{1t} and u_{2t} are incoherent at frequency 0 (cf. (6)).

Properties of the LSE and NBLSE were also studied by Robinson and Marinucci (2001, 2003) in case

$$\delta > 1/2, \tag{9}$$

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 used in the NBLSE 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 (2003) have established it for estimates asymptotically equivalent to the MLE, allowing for consistent estimation of unknown γ and δ and a vector θ 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 (2003) found the limit distribution unaffected by the question of whether θ, γ and δ are known or unknown.

The present paper focuses on the case of "weak fractional cointegration"

$$\beta \overset{def}{=} \delta - \gamma < 1/2, \tag{10}$$

where substantially different asymptotics prevail, impacting also on the question of how δ and γ should be estimated. Under (10), since $\Delta^\gamma y_t$ and $\Delta^\gamma x_t$ are

$I(\beta)$, they are asymptotically stationary, and one anticipates the existence of $\sqrt{n}$-consistent and asymptotically normal estimates of ν; the LSE and NBLSE converge slower than this owing to the dominance of bias due to (6). Under (10), the gain of a cointegration analysis is clearly less than when $\beta \geq 1/2$, for example in the $CI(1, 0)$ case. Nevertheless the identification of such structure is useful, and a variety of empirical evidence appears to support (10).

When cointegrated observables are stationary, and cointegrating errors are not antipersistent (so (4), (8) hold), (10) is inevitable. Andersen *et al* (2001) detected stationary long memory and co-movement in statistics derived from high-frequency transaction prices. Christensen and Prabhala (1998), Christensen and Nielsen (2004) found integration orders between 0.35 and 0.4 in implied and realized volatilities, and $I(0)$ cointegrating errors. In Robinson and Yajima's (2002) cointegration analysis of spot closing prices of crude oil, most estimated integration orders were less than 0.5. More generally, interest in the possibility of cointegration between stationary financial series is developing, and Robinson and Marinucci (2003) argued that it can be difficult to distinguish between a unit root process and some stationary long memory ones.

In other cases of (10), observables are nonstationary. When they have a unit root, so $\delta = 1$, it is implied that cointegrating errors are also nonstationary, albeit mean-reverting, in which case the cointegrating relation does not have the usual kind of "equilibrium" interpretation. Nevertheless a dimensionality reduction still occurs, empirical evidence for the phenomenon can be found, and the case $\gamma > 1/2$ has been stressed by Mármol and Velasco (2004). Diebold, Husted and Rush (1991) represented real exchange rates as errors in cointegrated, and apparently unit-root, nominal exchange rates and prices, and found them in some cases to be nonstationary. Similar mixed outcomes can be found in work of Cheung and Lai (1993) (investigating the PPP hypothesis), Baillie and Bollerslev (1994a) and Kim and Phillips (2000) (cointegration between spot exchange rates), Baillie and Bollerslev (1994b) (analyzing the forward premium) and Crato and Rothman (1994) (cointegration between exchange rates). On the other hand, there may be no strong reason to focus on $\delta = 1$ in a fractional context; autoregression-based unit root tests, such as those of Dickey and Fuller (1979), do not have good power against fractional alternatives, and though fractional-based tests have been developed (see e.g. Robinson, 1994b, Dolado, Gonzalo and Mayoral, 2002) one can treat δ as unknown. In this case, empirical evidence of $\delta > 1/2$ with $\beta < 1/2$ was found by Dueker and Startz (1998) (cointegration between US and Canadian bond rates) and Robinson and Marinucci (2003) (cointegration between stock prices and dividends, and between monetary aggregates), with estimates of δ variously less than and greater than 1.

Here we are principally concerned with estimating ν, under (10). Most of the empirical studies reported above employ semiparametric estimates of integration orders, with convergence rates slower than $\sqrt{n}$, so estimates of ν depending on them will, like the LSE and NBLSE, be less than $\sqrt{n}$-consistent. Achieving $\sqrt{n}$-consistency requires a parametric approach. Under both $\beta < 1/2$ and $\beta > 1/2$ the Gaussian MLE appears to have optimality properties and to provide Wald test statistics with null χ^2 limit distributions, and so should handle multivariate systems containing more than one cointegrating relation, where both $\beta < 1/2$ and $\beta > 1/2$ might occur. However, asymptotic properties of the MLE have yet to be developed, in case of autocorrelated u_t, and in (1) they can be achieved by a computationally simpler approach when $\beta > 1/2$, as described by Robinson and Hualde (2003), whereas this is not the case when $\beta < 1/2$.

To describe the theoretical background to inference when $\beta < 1/2$, note first that if γ and δ are known, while u_t is known to be white noise with unknown covariance matrix Ω, then the MLE of ν is given in closed form, and may be computed as an added-variable LSE, as pursued in the following section. When γ and/or δ are unknown, and u_t has parametric autocorrelation (such as following a vector autoregression (VAR)), then 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 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 (8), asymptotic 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 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 essential here (due to correlation between x_t and u_{1t}) that they be estimated $\sqrt{n}$-consistently. 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 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 a VAR, which we allow for. We propose $\sqrt{n}$-consistent and asymptotically normal estimates that require only univariate optimizations.

We mention finally other work on developing asymptotic inference on fractional cointegration, which employs a different definition of $I(d)$ processes for $d \neq 0$: for $v_t \sim I(0)$, we have $\Delta^{-d} v_t \sim I(d)$ for $|d| < 1/2$ and $\sum_{s=1}^{t} \Delta^{1-d} v_s \sim I(d)$ for $1/2 < d < 3/2$. This kind of fractional process has been called "Type I", and ours "Type II". Jeganathan (1999, 2001) considered such a "Type I" version of (1), for $|\gamma| < 1/2$ and $-1/2 < \delta < 3/2$, in a purely fractional context, such that v_t in the above definition is a white noise sequence. Assuming the distribution of the white noise inputs is of completely known (not necessarily Gaussian) form, Jeganathan (1999, 2001) considered limit distribution theory of optimal estimates of ν, including the mixed normal limit finding when $\beta > 1/2$, and Jeganathan (2001), using rates suggested by Robinson (2000), derived $\sqrt{n}$-consistency and asymptotic normality when $\beta < 1/2$; Jeganathan (2001) also covered the case $\beta = 1/2$. Though including some discussion of estimation of γ and δ, Jeganathan (1999, 2001) assumed them known in his theory. Aside from this, the Gaussian version of his estimates is the same as ours in case our u_t is white noise, though we do not assume Gaussianity. For the same, "Type I", definition of fractional processes, but with $I(0)$ inputs having nonparametric autocorrelation (implying a semiparametric model) Dolado and Mármol (1996), Kim and Phillips (2000) developed methods and theory in cases when $\beta > 1/2$.

The basic structure of our estimates of ν is described in the following section. Section 3 provides asymptotic theory in case γ and δ are known, with proofs and some technical details left to appendices. 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

Write (1) as

$$z_t(\gamma,\delta) = \zeta x_t(\gamma)\nu + u_t^{\#}, \tag{11}$$

where we introduce the notation

$$v_t(c) = \Delta^c v_t^{\#} \tag{12}$$

for a generic sequence v_t, and define

$$z_t(c,d) = (y_t(c), x_t(d))', \quad \zeta = (1,0)'. \tag{13}$$

We take u_t to be generated by the VAR

$$u_t = \sum_{j=1}^{p} B_j u_{t-j} + \varepsilon_t, \tag{14}$$

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 (14) to mean white noise u_t when $p=0$.

From (11) and (14) we have

$$z_t(\gamma,\delta) - \sum_{j=1}^{p} B_j z_{t-j}(\gamma,\delta) = \nu\left\{\zeta x_t(\gamma) - \sum_{j=1}^{p} B_j \zeta x_{t-j}(\gamma)\right\} + \varepsilon_t^{+}, \quad t\geq 1, \tag{15}$$

where

$$\varepsilon_t^{+} = u_1 1\,(t=1) + \left\{u_t - \sum_{j=1}^{t-1} B_j u_{t-j}\right\} 1\,(t=2,...,p) + \varepsilon_t 1\,(t>p)\,. \tag{16}$$

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 (15) can be written as

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

whence the first equation can be written as

$$y_t(\gamma) = \nu x_t(\gamma) + \rho x_t(\delta) + \sum_{j=1}^{p} \left(B_{1j} - \rho B_{2j}\right) z_{t-j}(\gamma,\delta) - \nu \sum_{j=1}^{p} \left(B_{1j} - \rho B_{2j}\right) \zeta x_{t-j}(\gamma) + \varepsilon_{1.2,t}, \tag{18}$$

where $\varepsilon_{1.2,t} = \varepsilon_{1t} - \rho\varepsilon_{2t}$, $\rho = E(\varepsilon_{1t}\varepsilon_{2t})/E(\varepsilon_{2t}^2)$; (18) 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 Q, 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 (18) as

$$y_t(\gamma) = \theta' Q Z_t(\gamma,\delta) + \varepsilon_{1.2,t}, \tag{19}$$

where

$$Z_t(c,d) = \left(x_t(c), x_t(d), w'_{t-1}(c,d), ..., w'_{t-p}(c,d)\right)', \tag{20}$$

$$w_t(c,d) = \left(x_t(c), x_t(d), y_t(c)\right)', \tag{21}$$

and the $q \times 1$ vector θ consists of coefficients that are not *a priori* zero, being (in some cases nonlinear) functions of ν, ρ, and the B_{ij}.

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

$$\widehat{\theta}(c,d) = G(c,d)^{-1} g(c,d), \tag{22}$$

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

$$G(c,d) = Q\frac{1}{n}\sum_{t=p+1}^{n} Z_t(c,d) Z'_t(c,d) Q', \quad g(c,d) = Q\frac{1}{n}\sum_{t=p+1}^{n} Z_t(c,d) y_t(c). \tag{23}$$

For example, in case $p = 1$, if u_{1t} is white noise while u_{2t} is AR(1), then $q = 3$ and (18) becomes

$$y_t(\gamma) = \nu x_t(\gamma) + \rho x_t(\delta) - \rho B_{221} x_{t-1}(\delta) + \varepsilon_{1.2,t}, \tag{24}$$

where B_{22j} is the second element of B_{2j}. Notice that ν, ρ and B_{221} are all identified in (24), but it is apparent from comparison of (18) with (19) 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

$$\widehat{\nu}(c,d) = 1' G(c,d)^{-1} g(c,d), \tag{25}$$

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

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 $\sqrt{n}$-consistent and computationally considerably simpler. Notice that over-specification of p results in a further efficiency loss, but under-specification 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 (14) 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 $\sqrt{n}$-consistency and asymptotic normality of $\widehat{\theta}(\gamma, \delta)$, and hence of $\widehat{\nu}(\gamma, \delta)$. We assume in addition to the description of (14) that the ε_t are stationary and ergodic with finite fourth moment, satisfying also

$$E\left(\left.\varepsilon_t\right|\mathcal{F}_{t-1}\right) = 0, \quad E\left(\left.\varepsilon_t\varepsilon_t'\right|\mathcal{F}_{t-1}\right) = \Omega \tag{26}$$

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. Noting from (1) that

$$x_t(\gamma) = \sum_{j=0}^{t-1} a_j(\beta)u_{2,t-j}, \quad t > 0; \quad = 0, \quad t \leq 0, \tag{27}$$

define θ

$$\overline{x}_t(\gamma) = \sum_{j=\max(t,0)}^{\infty} a_j(\beta)u_{2,t-j}, \quad \widetilde{x}_t(\gamma) = x_t(\gamma) + \overline{x}_t(\gamma), \tag{28}$$

so that because of (10), $\widetilde{x}_t(\gamma)$, $t = 0, \pm 1, ...$, is a covariance stationary sequence. Likewise, so is

$$\widetilde{y}_t(\gamma) = \nu\widetilde{x}_t(\gamma) + u_{1t}, \tag{29}$$

as is u_{2t}. Now define

$$\widetilde{w}_t = \left(\widetilde{x}_t(\gamma), u_{2t}, \widetilde{y}_t(\gamma)\right)', \quad \widetilde{Z}_t = \left(\widetilde{x}_t(\gamma), u_{2t}, \widetilde{w}_{t-1}', ..., \widetilde{w}_{t-p}'\right)', \tag{30}$$

$$\Phi = E(\widetilde{Z}_t\widetilde{Z}_t'), \quad \Psi = E\left(\varepsilon_{1.2,t}^2\widetilde{Z}_t\widetilde{Z}_t'\right). \tag{31}$$

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

Theorem 3.1 *Under (1), (4), (5), (10) and the conditions in the sentence containing (26), if γ and δ are known*

$$n^{1/2}\left\{\widehat{\theta}(\gamma, \delta) - \theta\right\} \to_d N\left(0, (Q\Phi Q')^{-1}Q\Psi Q'(Q\Phi Q')^{-1}\right), \tag{32}$$

as $n \to \infty$, and the covariance matrix on the right hand side is consistently estimated by

$$G(\gamma,\delta)^{-1}K(\gamma,\delta)G(\gamma,\delta)^{-1}, \tag{33}$$

where

$$K(c,d) = Q\frac{1}{n}\sum_{t=p+1}^{n}\widehat{\varepsilon}^2_{1.2,t}(c,d)Z_t(c,d)Z_t'(c,d)Q', \tag{34}$$

in which

$$\widehat{\varepsilon}_{1.2,t}(c,d) = y_t(c) - \widehat{\theta}(c,d)'QZ_t(c,d). \tag{35}$$

Remark 3.1 As anticipated, for $p \geq 1$, $\widehat{\nu}(\gamma,\delta)$ is inefficient relative to the Gaussian MLE, because it ignores the nonlinear restrictions on θ.

Remark 3.2 Over-parameterization in the B_j results in further loss of efficiency. 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 (24). Then if in fact u_{2t} is also white noise the limiting variance of $n^{1/2}\{\widehat{\nu}(\gamma,\delta)-\nu\}$ is

$$\omega^2_{1.2}/\{\omega^2_2\sum_{j=p+1}^{\infty}a^2_j(\beta)\}, \tag{36}$$

where $\omega^2_2 = E(\varepsilon^2_{2t})$, $\omega^2_{1.2} = E(\varepsilon^2_{1.2,t})$; (36) is increasing in p. As a simpler alternative to (33), (34), we can consistently estimate (36) by

$$\widehat{\omega}^2_{1.2}(\gamma,\delta)\left(1'G(\gamma,\delta)1\right)^{-1}, \tag{37}$$

where

$$\widehat{\omega}^2_{1.2}(\gamma,\delta) = \frac{1}{n}\sum_{t=p+1}^{n}\widehat{\varepsilon}^2_{1.2,t}(\gamma,\delta). \tag{38}$$

Note that (36) and (37) also apply in case $p=0$ is correctly taken in the estimation, when $\widehat{\nu}(\gamma,\delta)$ is equivalent to the Gaussian MLE, and (36) becomes

$$\omega^2_{1.2}/\left\{\frac{2^{-4\beta}\omega^2_2}{\pi}B\left(1/2-\beta,1/2-\beta\right)-1\right\}. \tag{39}$$

Note also that (36) and (39) do not depend on fourth cumulants of ε_t. If 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 (32)), in general depends on the fourth cumulant of $\varepsilon_{1.2,t}$, $\varepsilon_{1.2,t}$, ε_{2t} and ε_{2t}, though of course the latter is zero under Gaussianity.

Remark 3.3 On the other hand, under-parameterization of the B_j produces inconsistency of $\widehat{\nu}(\gamma,\delta)$, as when u_t is actually $AR(p+1)$. 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 3.2.

Remark 3.4 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. $\rho = 0$.

Remark 3.5 Though we assume (10) throughout, when in fact $\beta > 1/2$, $\hat{\nu}(\gamma, \delta)$ is as efficient as the Gaussian MLE. In particular, it can be shown to approximate the estimate of Robinson and Hualde (2003), which for $\beta > \frac{1}{2}$ has a limiting mixed normal distribution when the estimates of the parameters describing the short memory process u_t converge suitably fast, but need not themselves be asymptotically efficient.

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 might then be plugged into (25). Instead, we propose estimates of γ and δ that require two univariate nonlinear optimizations, in place of one bivariate one. The computational advantage in this would be intensified in extensions to systems involving a greater number of integration orders.

Write the second likelihood equation of (1) as

$$x_t(\delta) = \varepsilon_{2t}, \quad t \geq 1. \tag{40}$$

We estimate δ by

$$\widetilde{\delta}_0 = \arg\min_{d \in \mathcal{D}} S_0(d), \tag{41}$$

for a closed interval $\mathcal{D}$ and (cf. Beran, 1995),

$$S_0(d) = \sum_{t=1}^{n} x_t^2(d). \tag{42}$$

We then estimate γ by

$$\widetilde{\gamma}_0 = \arg\min_{c \in \mathcal{C}} T_0(c), \tag{43}$$

for a closed interval $\mathcal{C}$ (presumably a subset of $[0, \widetilde{\delta}_0]$) and

$$T_0(c) = \sum_{t=1}^{n} \left\{ y_t(c) - \widehat{\nu}(c, \widetilde{\delta}_0) x_t(c) - \widehat{\rho}(c, \widetilde{\delta}_0) x_t(\widetilde{\delta}_0) \right\}^2, \tag{44}$$

where $\widehat{\nu}(c, d)$ is given by (25), taking $p = 0$, and $\widehat{\rho}(c, d)$ is the second element of $\widehat{\theta}(c, d)$ in this case. The presence of c as argument in $y_t(c)$, and indeed of d in $x_t(d)$ of (42), 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

$$B_j \text{ is upper-triangular, } j = 1, ..., p. \tag{45}$$

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

$$x_t(\delta) - \phi' R X_t(\delta) = \varepsilon_{2t}, \tag{46}$$

with

$$X_t(d) = (x_{t-1}(d), ..., x_{t-p}(d))', \tag{47}$$

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 $r \times 1$ vector ϕ collecting the B_{22j} that are not *a priori* zero. The prescription (46) 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

$$\widehat{\phi}(d) = H(d)^{-1} h(d), \tag{48}$$

where

$$H(d) = R \frac{1}{n} \sum_{t=p+1}^{n} X_t(d) X_t'(d) R', \quad h(d) = R \frac{1}{n} \sum_{t=p+1}^{n} X_t(d) x_t(d). \tag{49}$$

First estimate δ by

$$\widetilde{\delta}_p = \arg\min_{d \in \mathcal{D}} S_p(d), \tag{50}$$

where

$$S_p(d) = \sum_{t=p+1}^{n} \left\{ x_t(d) - \widehat{\phi}(d)' R X_t(d) \right\}^2. \tag{51}$$

Then estimate γ by

$$\widetilde{\gamma}_p = \arg\min_{c \in \mathcal{C}} T_p(c), \tag{52}$$

where

$$T_p(c) = \sum_{t=p+1}^{n} \left\{ y_t(c) - \widehat{\theta}(c, \widetilde{\delta}_p)' Q Z_t(c, \widetilde{\delta}_p) \right\}^2. \tag{53}$$

As abbreviating notation, we denote throughout, for any $p \geq 0$, $\widetilde{\delta} = \widetilde{\delta}_p$, $\widetilde{\gamma} = \widetilde{\gamma}_p$. The proof of the following theorem is omitted for the following reasons. When $\delta < 1/2$ (and $\mathcal{D} \subset (0, 1/2)$) the proof of limit behaviour of $\widetilde{\delta}$ does not greatly differ from proofs of Fox and Taqqu (1986), Giraitis and Surgailis (1990); their x_t is actually stationary, not just asymptotically, and their objective functions differ from (42) and (51), though with $p = 0$ their estimates have equal asymptotic efficiency to our $\widetilde{\delta}_0$. When the possibility that $\delta > 1/2$ is allowed in the choice of $\mathcal{D}$, there is a difficulty in proving consistency of $\widetilde{\delta}$ if $\mathcal{D}$ includes $d \leq \delta - 1/2$, due to a lack of uniform convergence of $S_p(d)$ on $\mathcal{D}$. Since δ is unknown, there is no guarantee of avoiding this problem. Velasco and Robinson (2000) established consistency, and thence asymptotic normality with $\sqrt{n}$ rate, of an alternative estimate of δ allowing $\mathcal{D}$ to be arbitrarily large, but for "Type I" processes and employing tapering (which tends to inflate variance). Hualde and Robinson (2004) have recently done the same for $\widetilde{\delta}$ in our setting, with the unimportant difference that their linear process for x_t has scalar innovation, and is not nested in a non-diagonal bivariate system. In our setting, and whether or not $\mathcal{D} \subset (0, 1/2)$, the proof of Theorem 4.1 proceeds by establishing consistency of $\widetilde{\delta}$, following Hualde and Robinson (2004), then consistency of $\widetilde{\gamma}$, allowing for the extra complexity involved in working with residuals, and then employing the Cramer-Wold device and relatively straightforward and tedious arguments.

Theorem 4.1 *Under (1), (4), (5), (10), (45), the conditions in the sentence containing (26) and $\gamma \in \mathcal{C}$, $\delta \in \mathcal{D}$,*

$$n^{1/2} \begin{bmatrix} \widehat{\nu}(\widetilde{\gamma}, \widetilde{\delta}) - \nu \\ \widetilde{\gamma} - \gamma \\ \widetilde{\delta} - \delta \end{bmatrix} \to_d N(0, ABA'), \quad (54)$$

as $n \to \infty$, 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. 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 (10).

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. In fact, Robinson and Hualde (2003) showed that in this case the asymptotic distributions of $\widehat{\nu}(\gamma, \delta)$ and $\widehat{\nu}(\widetilde{\gamma}, \widetilde{\delta})$ are the same, due to $\widetilde{\gamma}$, $\widetilde{\delta}$ still being $\sqrt{n}$-consistent.

Remark 4.4 Robinson and Hualde (2003) suggest use of residuals from the LSE or NBLSE of ν in the estimation of γ when $\beta > 1/2$. However, the LSE and NBLSE are less-than-$\sqrt{n}$-consistent under (10), and so it appears that the resulting estimates of γ will not achieve the $\sqrt{n}$-consistency needed to provide a $\sqrt{n}$-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. (41)), whereas the first equation also contains relative information. However, the estimates can be updated to efficiency by a single Newton step.

Remark 4.6 The paper has taken existence of cointegration, and $\beta < 1/2$, for granted. In practice these properties will have to be established, and our estimation of ν will form the final step. Some discussion of methodology has already appeared in Marinucci and Robinson (2001), Robinson and Yajima (2002) and Robinson and Marinucci (2003). This has stressed a semiparametric approach, recognizing that a parametric model for u_t (i.e. knowledge of p in our VAR case) is unlikely to be known *a priori*. A natural starting point is to test the necessary requirement of equality of integration orders of x_t and y_t. The literature on asymptotic inference for multivariate fractional models is rather small, and some of it assumes lack of cointegration, but the approaches of Robinson and Yajima (2002) and Hualde (2002) are available. Given a positive outcome, a test for existence of cointegration, such as those of Marinucci and Robinson (2001), Robinson and Yajima (2002), Mármol and Velasco (2004) can be conducted. Given a positive outcome, one can reject $\beta = 1/2$ against the alternative $\beta < 1/2$ if a suitably standardized $\bar{\delta} - \bar{\gamma} - 1/2$ is significantly negative relative to the standard normal distribution, $\bar{\delta}$ and $\bar{\gamma}$ being semiparametric estimates of δ and γ, employing residuals $y_t - \bar{\nu} x_t$ based on an initial consistent $\bar{\nu}$, such as the NBLSE. Then, using proxies $y_t(\bar{\gamma}) - \bar{\nu} x_t(\bar{\gamma})$ and $x_t(\bar{\gamma})$ for u_{1t} and u_{2t} respectively, the AR order of u_t can be identified, for example as described in the empirical examples of Section 6.

5. Monte Carlo Evidence

With the main aim of investigating the finite sample performance of our estimates of ν and associated rules of inference, a Monte Carlo experiment was carried out. There are two parts to our investigation, the first comparing our proposed estimates (with known and unknown integration orders) with the simplest one, the LSE, and the second evaluating in a simple framework the inefficiency of $\widehat{\nu}(\gamma, \delta)$ mentioned in Remark 3.1 with respect to two asymptotically efficient estimates of ν. In data generation from (1), (14), we took $p = 1$ throughout, with

$$B_1 = diag\left\{b_1, b_2\right\}, \tag{55}$$

where each b_i took values 0, 0.5, 0.9. The case $b_1 = b_2 = 0$ actually corresponds to $p = 0$ in (14), where u_t is a white noise vector. Likewise, $b_1 = 0$, $b_2 \neq 0$ corresponds to (24). We have employed in (55) abbreviating notation compared to (24), so $b_2 = B_{221}$. The ε_t in (14) 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:

$$(\gamma, \delta) = (0, 0.4)\,,\ (0.2, 0.4)\,,\ (0.4, 0.8)\,,\ (0.7, 1)\,, \tag{56}$$

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 (27), 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, in the first part of the experiment we computed estimates of the following three types:
(i) The LSE,

$$\overline{\nu}_0 = \sum_{t=1}^{n} x_t y_t / \sum_{t=1}^{n} x_t^2. \tag{57}$$

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

By "correct specification" we mean that all prior zero restrictions on B_1 in (55), 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 $Z_t(c,d) = \left(x_t(c), x_t(d), w_{t-1}'(c,d)\right)'$. Knowledge of $\rho = 0$ was never used. Table 1 records convergence rates of the LSE and, under the heading "optimal", of $\overline{\nu}_I$, $\overline{\nu}_F$.

(Table 1 about here)

We now describe how $\widetilde{\delta}$ and $\widetilde{\gamma}$ in $\overline{\nu}_F$ were computed. In estimating δ, we fixed $\mathcal{D} = [\delta - 1, \delta + 1]$ in (50). A $\mathcal{D}$ of length 2 may often be adequate. In estimating γ, we fixed $\mathcal{C} = [\widetilde{\delta} - 2.05, \widetilde{\delta} - 0.05]$ in (52). The upper bound seems reasonable since a very small β is unlikely to be detectable, indeed there is then near loss of identifiability and very poor behaviour of estimates of ν.

The estimates $\overline{\nu}_I$, $\overline{\nu}_F$ (but not $\overline{\nu}_0$) are invariant to (γ, δ) combinations with the same β, provided the fractionally integrated processes are generated from the same u_t sequence. Thus we do not report results for $(\gamma, \delta) = (0.4, 0.8)$ in tables where only $\overline{\nu}_I$ and $\overline{\nu}_F$ are involved. Similarly, $\widetilde{\delta} - \delta$ is invariant to δ, so the reported results apply to any δ, whereas $\widetilde{\gamma} - \gamma$ is invariant to (γ, δ) combinations with the same β, so we again omit results for $(\gamma, \delta) = (0.4, 0.8)$.

(Tables 2-7 about here)

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 (b_1, b_2) combinations covered in the experiment are included, in order to conserve on 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 relative performance of $\overline{\nu}_0$, $\overline{\nu}_I$ and $\overline{\nu}_F$ mentioned above is maintained 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 $b_1 = b_2 = 0.9$, $\beta = 0.4$ 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). Focussing 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 $\beta = 0.4$. 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 $\overline{\nu}_F$, except for the case $b_1 = 0.9$, $b_2 = 0$. $\overline{\nu}_I$ 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, $(0.2, 0.4)$ being clearly the worst case, though again bias decreases with n. As anticipated, over-specification (Table 7) makes little difference to $\overline{\nu}_I$, which does much better than $\overline{\nu}_0$, but $\overline{\nu}_F$ is damaged (especially for $\beta = 0.4$) by poor estimates of the integration orders. However, small reductions on the optimizing intervals $\mathcal{C}$, $\mathcal{D}$, cause very significant improvements in $\overline{\nu}_F$ (and in fact in the estimates of γ, δ).

(Tables 8-11 about here)

Tables 8-11 contain Monte Carlo standard deviations (SD) 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 focussing on only few cases. In fact, $\overline{\nu}_0$ was superior to $\overline{\nu}_I$ for most combinations, including those not displayed, with $\overline{\nu}_F$ a poor third. With correct specification, this was most notably evident 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 SD of $\overline{\nu}_F$ relative to $\overline{\nu}_I$. 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 SD than $\overline{\nu}_0$. Nevertheless, it must be borne in mind that the paper's motivation is not to minimize variance but rather to achieve $\sqrt{n}$-consistency and asymptotic normality in a fairly general context, which the LSE $\overline{\nu}_0$ does not provide.

We now examine the usefulness of the limit distributional properties of $\overline{\nu}_I$ and $\overline{\nu}_F$ by examining the size of Wald tests. We computed

$$W_I = \frac{(\overline{\nu}_I - \nu)^2 n}{[G(\gamma,\delta)^{-1}K(\gamma,\delta)G(\gamma,\delta)^{-1}]_{(1)}},\ W_F = \frac{(\overline{\nu}_F - \nu)^2 n}{\left[\widehat{A}\widehat{B}\widehat{A}'\right]_{(1)}}, \tag{58}$$

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.

(Tables 12-17 about here)

With correct specification, even for $b_1 = b_2 = 0$ (Table 12), sizes of the infeasible statistic W_I are somewhat too large, and autocorrelation in u_t exacerbates 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 more variation across parameter values, and they are sometimes actually less than nominal values. 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 sometimes the results are extraordinarily good. As expected, the effect of mis-specification is more dramatic than for W_I. With over-specification (Table 17), sizes are mainly less than nominal values, but in general approximate them as n increases. 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 an empirical analysis of fractional cointegration, and so we also give some space to the performance of $\widetilde{\delta}$ and $\widetilde{\gamma}$, and to Wald tests for δ and γ based on Theorem 4.1.

(Tables 18 and 19 about here)

Tables 18 and 19 report Monte Carlo bias and SD of $\widetilde{\delta}$ for the same values of 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 (42), (51)), 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.

Biases from $S_0(d)$ with $b_2 = 0$ look satisfactory even for $n = 64$, and decrease in n. For $S_1(d)$ with $b_2 = 0.5, 0.9$, there is some deterioration, but performance is still acceptable. For $S_1(d)$ with $b_2 = 0$ results are worse, but small reductions in $\mathcal{D}$ have a large positive impact on $\widetilde{\delta}$. In this case the negative bias of $\widetilde{\delta}$ is somehow expected, as the estimated (non-existent) AR component in u_{2t} accounts for some of the autocorrelation structure. Unsurprisingly, there is severe bias, increasing with b_2, when $S_0(d)$ is used with $b_2 \neq 0$. SD in the correctly specified and over-specified cases is, as expected, worse for AR u_t.

(Tables 20 and 21 about here)

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

$$W_\delta = \frac{(\widetilde{\delta} - \delta)^2 n}{\left[\widehat{A}\widehat{B}\widehat{A}'\right]_{(3)}}, \tag{59}$$

based on Theorem 4.1, with respect to nominal sizes $\alpha = 0.05, 0.1$ respectively. As expected, under mis-specification they are far too large. Otherwise, while still too large (especially for over-specification) in some cases they are not bad, and decrease in n, ones for $\alpha = 0.1$ being best.

(Tables 22-25 about here)

Tables 22-25 give corresponding results for $\widetilde{\gamma}$, with $b_1 = b_2 = b$ taking values 0, 0.5, 0.9. Our estimation procedure being sequential, we consider two categories, $S_0(d)$ followed by $T_0(c)$ (44), and $S_1(d)$ followed by $T_1(c)$ (53), 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 SD results of Tables 22 and 23 exhibit some variation across (γ, δ), and surprisingly biases are much less for $b = 0.9$ than for $b = 0.5$, possibly due to cancellation. For the Wald statistic

$$W_\gamma = \frac{(\widetilde{\gamma} - \gamma)^2 n}{\left[\widehat{A}\widehat{B}\widehat{A}'\right]_{(2)}}, \tag{60}$$

more size variation is also found, in Tables 24 and 25, than for W_δ, with results for $b = 0.9$ being substantially better than for other cases under correct specification.

(Table 26 about here)

For the second part of the study, we focus on a situation where it is straightforward to derive asymptotically efficient estimates of ν, and we compare their Monte Carlo variance with that of $\overline{\nu}_I$. We consider only the case where in (55), $b_1 = 0.5, 0.9$, $b_2 = 0$. The first efficient estimate we calculate is the Gaussian MLE with known b_1, which, in view of (18) is identical to the LSE of ν in the equation

$$y_t(\gamma) - b_1 y_{t-1}(\gamma) = \nu(x_t(\gamma) - b_1 x_{t-1}(\gamma)) + \rho x_t(\delta) + \varepsilon_{1.2t}. \tag{61}$$

We also consider a two-stage approach where in the first step we estimate b_1 by

$$\widehat{b}_1 = \frac{\sum_{t=2}^{n} \widehat{u}_{1t}\widehat{u}_{1,t-1}}{\sum_{t=2}^{n} \widehat{u}_{1,t-1}^2},\ \widehat{u}_{1t} = y_t(\gamma) - \overline{\nu}_I x_t(\gamma), \tag{62}$$

and in the second compute the estimate of ν as in the infeasible situation but replacing b_1 by $\widehat{b}_1$. We report in Table 26 the "efficiency ratios" r_1 and r_2, which are the Monte Carlo variance of $\overline{\nu}_I$ divided by either that of the Gaussian MLE with known b_1 (r_1) or that of the feasible estimate (r_2). Note that r_1 and r_2 are invariant to the value of $E(\varepsilon_{2t}^2)$, provided the estimates are computed from the same u_t sequence. In general, results are little affected by changes in ρ, and the loss of efficiency of $\overline{\nu}_I$ is larger for smaller β and larger b_1. As expected, $\overline{\nu}_I$ is more inefficient relative to the infeasible MLE, and this is accentuated the larger and smaller b_1 and β are respectively. In the comparison with the infeasible MLE, the efficiency loss is reduced as n increases, the reverse happening for the feasible estimate. On the limited evidence provided by our simple experiment, it seems worth improving efficiency by incorporating restrictions on θ. Undoubtedly more iterations could further improve matters.

6. Empirical Examples

Using a methodology involving the LSE and NBLSE of ν, and semiparametric estimates of ν, Robinson and Marinucci (2003) 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 investigated 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 (14) 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 (57) (and computed by Robinson and Marinucci, 2003). To estimate γ and δ, we used semiparametric estimates (already computed by Marinucci and Robinson, 2001, Robinson and Marinucci, 2003) 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). Using 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 of memory parameters based on the x_t/residuals provided by Marinucci and Robinson (2001), Robinson and Marinucci (2003). When this led to contradictory models for the u_{it} the analysis was continued with both.

We also took this opportunity to examine the matter of truncation, which in one form or another always arises with fractional models, and perhaps most acutely when nonstationary data are involved. When estimated innovations from a stationary fractional model are computed, the (infinite) AR representation has to be truncated because the data begin at time "1", not at time "$-\infty$". 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 begin, and if we were to drop the first observation, say, and start at the next one, the degree of filtering applied to all subsequent observations would change, and this could have a marked effect, especially with nonstationary data, even though filtering is here applied after de-meaning. 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 (2003), 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 (14), the other with $p = 1$ in (14) with $B_1 = diag\,(b_1, 0)$.

(Table 27 about here)

In case u_{1t} and u_{2t} are both white noise, Table 27 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 (18) for $p=0$ with $\widetilde{\gamma}$, $\widetilde{\delta}$, replacing γ, δ, and the correlation $Corr\,(\varepsilon_{1t},\varepsilon_{2t})$ is estimated by

$$r=\widehat{\rho}(\widetilde{\gamma},\widetilde{\delta})(\widehat{\sigma}_{22}/\widehat{\sigma}_{11})^{\frac{1}{2}}, \tag{63}$$

where

$$\widehat{\sigma}_{11}=n^{-1}\sum\nolimits_t'\left(y_t(\widetilde{\gamma})-\widehat{\nu}(\widetilde{\gamma},\widetilde{\delta})x_t(\widetilde{\gamma})\right)^2,\ \widehat{\sigma}_{22}=n^{-1}\sum\nolimits_t' x_t^2(\widetilde{\delta}), \tag{64}$$

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. Dropping the first observation does not affect $\widetilde{\delta}$, since $x_1\,(d)=x_1$ for any d.

(Table 28 about here)

The analysis with AR(1) u_{1t} in Table 28 presents a very different picture. Here, we also report $\widehat{b}_1$ and $\widehat{\nu b_1}$, which are estimated coefficients of $y_{t-1}(\widetilde{\gamma})$ and $-x_{t-1}(\widetilde{\gamma})$ in the regression (cf. (18)) 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 supposed to be white noise, the estimates of δ are identical to those in Table 27, but those 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 Section 1, it would require much heavier computation. Notice that $\widehat{\nu b_1}$ looks quite consistent with $\widehat{\nu}$ 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.

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 (2001) 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 (14). The results are reported in Table 29. 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.

(Tables 29 and 30 about here)

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 Marinucci and Robinson's (2001) and Robinson and Marinucci's (2003) analyses again suggested 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 30, 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.

Appendix A: Proof of Theorem 3.1

We prove first that Φ is nonsingular, which ensures existence of the inverses in (32). 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 (29) 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 a $1 \times (p+1)$ vector of zeros. Define by $f(\lambda)$ the spectral density matrix of u_t, and note from positive definiteness 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

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

by a non-negative definite matrix, where 0, C and D 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 $C - DD'$ is positive definite. But for a $(p+1) \times 1$ vector $\zeta = (\zeta_i)$,

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

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

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

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

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

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

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

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

Further,

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

and so

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

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 (32). The proof of the final statement of the theorem is omitted as it is standard given (32) 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

$$\begin{aligned} \hat{a}_1' &= 1'\tilde{G}^{-1}, \quad \hat{a}_2 = -1'\tilde{\theta}_c \tilde{s}_{cc}^{-1}, & \text{(B.2)} \\ \hat{a}_3 &= 1'\tilde{\theta}_c \tilde{s}_{cc}^{-1} \tilde{s}_{cd} \tilde{s}_{dd}^{-1} - 1'\tilde{\theta}_d \tilde{s}_{dd}^{-1}, \quad \hat{a}_4 = -\tilde{s}_{cc}^{-1}, & \text{(B.3)} \\ \hat{a}_5 &= \tilde{s}_{cc}^{-1} \tilde{s}_{cd} \tilde{s}_{dd}^{-1}, \quad \hat{a}_6 = -\tilde{s}_{dd}^{-1}, & \text{(B.4)} \end{aligned}$$

in which

$$\begin{aligned} \tilde{\theta}_c &= \tilde{G}^{-1}\left(\tilde{g}_c - \tilde{G}_c\tilde{\theta}\right), \quad \tilde{\theta}_d = \tilde{G}^{-1}\left(\tilde{g}_d - \tilde{G}_d\tilde{\theta}\right), & \text{(B.5)} \\ \tilde{g}_c &= Q\frac{1}{n}\sum{}' \left\{Z_{tc}(\tilde{\gamma})y_t(\tilde{\gamma}) + Z_t(\tilde{\gamma}, \tilde{\delta})y_{tc}(\tilde{\gamma})\right\}, & \text{(B.6)} \\ \tilde{G}_c &= Q\frac{1}{n}\sum{}' \left\{Z_{tc}(\tilde{\gamma})Z_t'(\tilde{\gamma}, \tilde{\delta}) + Z_t(\tilde{\gamma}, \tilde{\delta})Z_{tc}'(\tilde{\gamma})\right\} Q', & \text{(B.7)} \\ \tilde{g}_d &= Q\frac{1}{n}\sum{}' Z_{td}(\tilde{\delta})y_t(\tilde{\gamma}), & \text{(B.8)} \\ \tilde{G}_d &= Q\frac{1}{n}\sum{}' \left\{Z_{td}(\tilde{\delta})Z_t'(\tilde{\gamma}, \tilde{\delta}) + Z_t(\tilde{\gamma}, \tilde{\delta})Z_{td}'(\tilde{\delta})\right\} Q', & \text{(B.9)} \end{aligned}$$

with

$$\begin{aligned} y_{tc}(\tilde{\gamma}) &= \log(1-L)y_t(\tilde{\gamma}), & \text{(B.10)} \\ Z_{tc}(\tilde{\gamma}) &= \log(1-L)\left\{x_t(\tilde{\gamma}), 0, x_{t-1}(\tilde{\gamma}), 0, y_{t-1}(\tilde{\gamma}), ..., x_{t-p}(\tilde{\gamma}), 0, y_{t-p}(\tilde{\gamma})\right\}', & \text{(B.11)} \\ Z_{td}(\widetilde{\delta}) &= \log(1-L)\left\{0, x_t(\widetilde{\delta}), 0, x_{t-1}(\widetilde{\delta}), 0, ..., 0, x_{t-p}(\widetilde{\delta}), 0\right\}', & \text{(B.12)} \end{aligned}$$

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

$$\begin{aligned} \tilde{v}_{tc} &= y_{tc}(\tilde{\gamma}) - \tilde{\theta}_c' QZ_t(\tilde{\gamma}, \tilde{\delta}) - \tilde{\theta}' QZ_{tc}(\tilde{\gamma}), & \text{(B.14)} \\ \tilde{v}_{td} &= -\tilde{\theta}_d' QZ_t(\tilde{\gamma}, \tilde{\delta}) - \tilde{\theta}' QZ_{td}(\tilde{\delta}), & \text{(B.15)} \\ \tilde{w}_{td} &= x_{td}(\tilde{\delta}) - \tilde{\phi}_d' RX_t(\tilde{\delta}) - \tilde{\phi}' RX_{td}(\tilde{\delta}), & \text{(B.16)} \\ x_{td}(\tilde{\delta}) &= \log(1-L)x_t(\tilde{\delta}), & \text{(B.17)} \\ X_{td}(\tilde{\delta}) &= \log(1-L)X_t(\tilde{d}), & \text{(B.18)} \\ \tilde{\phi}_d &= \tilde{H}^{-1}(\tilde{h}_d - \tilde{H}_d\tilde{\phi}), & \text{(B.19)} \\ \tilde{h}_d &= R\frac{1}{n}\sum{}' \left\{X_{td}(\tilde{\delta})x_t(\tilde{\delta}) + X_t(\tilde{\delta})x_{td}(\tilde{\delta})\right\}, & \text{(B.20)} \\ \tilde{H}_d &= R\frac{1}{n}\sum{}' \left\{X_{td}(\tilde{\delta})X_t'(\tilde{\delta}) + X_t(\tilde{\delta})X_{td}'(\tilde{\delta})\right\} R'. & \text{(B.21)} \end{aligned}$$

We also have

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

where

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

Acknowledgements

This research was supported by ESRC Grants R000238212 and R000239936. 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. We are grateful for the comments of a referee which have led to improvements in the paper.

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens H., 2001. The distribution of realized stock return volatility. Journal of Financial Economics 61, 43-76.

Baillie R.T., Bollerslev T., 1994a. Cointegration, fractional cointegration and exchange rate dynamics. The Journal of Finance 49 (2), 737-745.

Baillie R.T., Bollerslev T., 1994b. The long memory of the forward premium. Journal of International Money and Finance 13 (5), 565-571.

Beran, J., 1995. Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. Journal of the Royal Statistical Society, Ser. B 57, 659-672.

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.

Cheung, Y.W., Lai, K.S., 1993. A fractional cointegration analysis of purchasing power parity. Journal of Business and Economic Statistics 11, 103-112.

Christensen, B.J., Prabhala, N.R., 1998. The relation between implied and realized volatility. Journal of Financial Economics 50, 125-150.

Christensen, B.J., Nielsen, M. O., 2004. Semiparametric analysis of stationary fractional cointegration and the implied-realized volatility relation. Journal of Econometrics, forthcoming.

Crato, N., Rothman, P., 1994. A reappraisal of parity reversion for UK real exchange rates. Applied Economics Letters 1, 139-141.

Dickey, D.A., Fuller, W.A., 1979. Distribution of estimators of autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427-431.

Diebold, F.X., Husted, S., Rush, M., 1991. Real exchange rates under the gold standard. Journal of Political Economy 99, 1252-1271.

Dolado, J., Gonzalo, J., Mayoral, L., 2002. A fractional Dickey-Fuller test. Econometrica 70, 1963-2006.

Dolado, J., Mármol, F., 1996. Efficient estimation of cointegrating relationships among higher order and fractionally integrated processes. Banco de España-Servicio de Estudios, Documento de Trabajo 9617.

Dueker, M., Startz, R., 1998. Maximum-likelihood estimation of fractional cointegration with an application to US and Canadian bond rates. The Review of Economics and Statistics 80, 420-426.

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.

Giraitis, L., Surgailis, D., 1990. A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. Probability Theory and Related Fields 86, 87-104.

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

Hualde, J., 2002. Testing for the equality of orders of integration. Preprint.

Hualde, J., Robinson, P.M., 2004. Gaussian pseudo-maximum likelihood estimation of fractional time series models. Preprint.

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.

Mármol, F., Velasco, C., 2004. Consistent testing of cointegrating relationships. Econometrica 72, 1809-1844.

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

Robinson, P.M., 1994b. Efficient tests of nonstationary hypotheses. Journal of the American Statistical Association 89, 1420-1437.

Robinson, P.M., 1995a. Log-periodogram regression of time series with long range dependence. Annals of Statistics 23, 1048-1072.

Robinson, P.M., 1995b. Gaussian semiparametric estimation of long-range dependence. Annals of Statistics 23 1630-1661.

Robinson, P.M., 2000. Private communication.

Robinson, P.M., Hualde, J., 2003. Cointegration in fractional systems with unknown integration orders. Econometrica 71, 1727-1766.

Robinson, P.M., Marinucci, D., 2001. Narrow-band analysis of nonstationary processes. Annals of Statistics 29, 947-986.

Robinson, P.M., Marinucci, D., 2003. Semiparametric frequency-domain analysis of fractional cointegration. In Time Series with Long Memory, ed. by P. M. Robinson (Oxford: Oxford University Press, pp. 334-373).

Robinson, P. M., Yajima, Y., 2002. Determination of cointegrating rank in fractional systems. Journal of Econometrics 106, 217-241.

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 1

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, $\rho = 0$	$n^{.5}$	$n^{.5}$	$n^{.4}$	$n^{.3}$

TABLE 2

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

		n	64			128			256		
ρ	γ	δ	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$
-.5	0	.4	.000	.061	-.338	-.002	.017	-.320	-.003	.008	-.307
	.2	.4	.000	.124	-.401	-.005	.077	-.387	-.010	.048	-.377
	.4	.8	.000	.061	-.193	-.002	.017	-.151	-.003	.008	-.120
	.7	1	.000	.101	-.220	-.003	.040	-.176	-.006	.020	-.142
0	0	.4	-.006	-.006	-.007	-.001	.000	-.003	-.001	-.001	.000
	.2	.4	-.014	-.038	-.011	.000	-.001	-.005	-.003	-.007	.000
	.4	.8	-.006	-.006	-.015	-.001	.000	-.009	-.001	-.001	-.002
	.7	1	-.009	-.020	-.031	.000	.000	-.023	-.002	-.003	-.005
.5	0	.4	.001	-.089	.337	.005	-.016	.320	.003	-.009	.308
	.2	.4	-.001	-.179	.394	.009	-.081	.384	.006	-.056	.376
	.4	.8	.001	-.089	.192	.005	-.016	.155	.003	-.009	.120
	.7	1	.000	-.142	.214	.006	-.043	.182	.004	-.025	.143
.75	0	.4	.002	-.123	.511	.003	-.029	.481	.002	-.010	.460
	.2	.4	.003	-.212	.599	.007	-.125	.578	.006	-.077	.562
	.4	.8	.002	-.123	.287	.003	-.029	.226	.002	-.010	.176
	.7	1	.003	-.194	.315	.005	-.073	.258	.004	-.028	.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	-.015	.165	-.161	-.003	.102	-.136	-.005	.078	-.120
	.2	.4	-.041	.121	-.293	-.008	.096	-.266	-.006	.120	-.248
-.5	.4	.8	-.015	.165	-.147	-.003	.102	-.113	-.005	.078	-.088
	.7	1	-.024	.190	-.207	-.005	.080	-.166	-.006	.134	-.131
	0	.4	-.026	-.086	-.014	-.016	-.039	-.005	-.008	-.005	.000
	.2	.4	-.057	-.155	-.027	-.033	-.098	-.012	-.009	-.010	-.001
0	.4	.8	-.026	-.086	-.025	-.016	-.039	-.014	-.008	-.005	-.003
	.7	1	-.036	.036	-.043	-.022	-.093	-.030	-.008	.002	-.006
	0	.4	.016	-.208	.158	.004	-.145	.137	.005	-.073	.120
	.2	.4	.028	-.118	.281	.010	-.168	.267	.008	-.122	.247
.5	.4	.8	.016	-.208	.140	.004	-.145	.116	.005	-.073	.090
	.7	1	.019	-.269	.195	.006	-.144	.170	.006	-.081	.134
	0	.4	.027	-.278	.237	.010	-.149	.202	.007	-.068	.176
	.2	.4	.047	-.092	.421	.020	-.143	.390	.010	-.158	.364
.75	.4	.8	.027	-.278	.206	.010	-.149	.165	.007	-.068	.129
	.7	1	.034	-.278	.283	.013	-.215	.236	.008	-.139	.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	.036	-.142	.000	.044	-.128	.000	.032	-.119
	.2	.4	-.002	.058	-.203	.001	.065	-.189	-.001	.058	-.181
-.5	.4	.8	-.001	.036	-.083	.000	.044	-.065	.000	.032	-.052
	.7	1	-.001	.057	-.106	.000	.065	-.085	.000	.050	-.069
	0	.4	-.001	-.003	-.004	.001	.001	-.001	.001	.000	.000
	.2	.4	.001	-.018	-.008	.004	-.007	-.003	.003	.006	.000
0	.4	.8	-.001	-.003	-.008	.001	.001	-.005	.001	.000	-.001
	.7	1	.000	-.012	-.017	.002	.000	-.012	.002	.004	-.002
	0	.4	.006	-.034	.142	.004	-.030	.129	.001	-.029	.119
	.2	.4	.016	-.037	.201	.010	-.045	.189	.004	-.044	.180
.5	.4	.8	.006	-.034	.082	.004	-.030	.067	.001	-.029	.052
	.7	1	.009	-.048	.102	.006	-.053	.088	.002	-.041	.069
	0	.4	.004	-.061	.216	.002	-.059	.192	.000	-.047	.178
	.2	.4	.011	-.089	.305	.006	-.094	.283	.001	-.085	.269
.75	.4	.8	.004	-.061	.123	.002	-.059	.097	.000	-.047	.076
	.7	1	.006	-.093	.151	.003	-.091	.124	.001	-.073	.100

TABLE 5

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

		n	64			128			256		
ρ	γ	δ	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$
-.5	0	.4	-.118	-.189	-.758	-.040	-.099	-.755	-.014	-.046	-.746
	.2	.4	-.264	-.323	-1.05	-.110	-.155	-1.11	-.054	-.084	-1.14
	.4	.8	-.118	-.189	-1.05	-.040	-.099	-.965	-.014	-.046	-.852
	.7	1	-.172	-.329	-1.51	-.066	-.142	-1.41	-.029	-.064	-1.26
0	0	.4	.006	.006	-.039	.005	.042	-.015	.005	.011	-.002
	.2	.4	-.002	-.063	-.065	.009	.019	-.030	.005	.000	-.005
	.4	.8	.006	.006	-.119	.005	.042	-.082	.005	.011	-.013
	.7	1	.003	-.073	-.251	.006	.029	-.210	.005	.008	-.038
.5	0	.4	.129	.111	.714	.052	.124	.740	.018	.038	.741
	.2	.4	.258	.189	.970	.126	.167	1.07	.056	.067	1.12
	.4	.8	.129	.111	.994	.052	.124	.981	.018	.038	.854
	.7	1	.177	.173	1.42	.079	.147	1.46	.032	.052	1.27
.75	0	.4	.167	.190	1.09	.065	.128	1.11	.022	.039	1.11
	.2	.4	.363	.300	1.48	.172	.237	1.61	.079	.074	1.68
	.4	.8	.167	.190	1.48	.065	.128	1.42	.022	.039	1.25
	.7	1	.242	.291	2.08	.106	.197	2.05	.043	.058	1.83

TABLE 6

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

		n	64			128			256		
ρ	γ	δ	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$
-.5	0	.4	.000	1.13	-.242	-.001	.981	-.221	-.003	.851	-.208
	.2	.4	.000	1.95	-.346	-.003	1.84	-.328	-.009	1.64	-.316
	.4	.8	.000	1.13	-.167	-.001	.981	-.132	-.003	.851	-.105
	.7	1	.000	1.51	-.212	-.002	1.30	-.170	-.005	1.10	-.138
0	0	.4	-.005	1.79	-.008	.000	1.59	-.003	.000	1.40	.000
	.2	.4	-.010	3.40	-.016	.002	3.30	-.006	-.001	3.06	-.001
	.4	.8	-.005	1.79	-.017	.000	1.59	-.010	.000	1.40	-.002
	.7	1	-.007	2.50	-.033	.000	2.20	-.024	.000	1.91	-.005
.5	0	.4	.004	2.45	.240	.006	2.10	.222	.003	1.98	.208
	.2	.4	.008	5.04	.337	.013	4.73	.326	.007	4.45	.314
	.4	.8	.004	2.45	.164	.006	2.10	135	.003	1.98	.105
	.7	1	.005	3.54	.204	.008	3.03	177	.004	2.78	.140
.75	0	.4	.004	2.71	.365	.003	2.33	.332	.002	2.21	.310
	.2	.4	.009	5.88	.513	.008	5.46	.487	.006	5.04	.469
	.4	.8	.004	2.71	.244	.003	2.33	.196	.002	2.21	.154
	.7	1	.006	3.97	.300	.005	3.40	.250	.003	3.12	.201

TABLE 7

MONTE CARLO BIAS, $b_1 = b_2 = 0$, over-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	.020	.207	-.338	.013	.248	-.320	.021	.149	-.307
	.2	.4	.065	.012	-.401	.042	.042	-.387	.035	-.017	-.377
-.5	.4	.8	.020	.207	-.193	.013	.248	-.151	.021	.149	-.120
	.7	1	.035	.179	-.220	.022	.205	-.176	.026	.124	-.142
	0	.4	-.032	-.056	-.007	-.006	-.059	-.003	.007	-.140	.000
	.2	.4	-.036	-.058	-.011	.023	-.083	-.005	.027	-.108	.000
0	.4	.8	-.032	-.056	-.015	-.006	-.059	-.009	.007	-.140	-.002
	.7	1	-.034	-.043	-.031	.003	-.101	-.023	.014	-.115	-.005
	0	.4	.006	-.291	.337	.017	-.323	.320	-.005	-.290	.308
	.2	.4	.021	-.092	.394	.061	-.151	.384	.004	-.155	.376
.5	.4	.8	.006	-.291	.192	.017	-.323	.155	-.005	-.290	.120
	.7	1	.012	-.259	.214	.032	-.288	.182	-.001	-.263	.143
	0	.4	-.018	-.288	.511	.002	-.319	.481	-.016	-.187	.460
	.2	.4	-.034	-.103	.599	.016	-.102	.578	-.021	-.178	.562
.75	.4	.8	-.018	-.288	.287	.002	-.319	.226	-.016	-.187	.176
	.7	1	-.023	-.191	.315	.007	-.319	.258	-.017	-.210	.206

TABLE 8

MONTE CARLO S.D., $b_1 = b_2 = 0$, 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	.178	.429	.109	.112	.169	.084	.076	.104	.065
	.2	.4	.419	.788	.131	.274	.470	.102	.193	.318	.077
-.5	.4	.8	.178	.429	.154	.112	.169	.122	.076	.104	.092
	.7	1	.259	.583	.270	.167	.286	.237	.116	.183	.188
	0	.4	.212	.321	.107	.128	.151	.073	.086	.093	.049
	.2	.4	.489	.817	.141	.310	.469	.105	.217	.284	.076
0	.4	.8	.212	.321	.171	.128	.151	.128	.086	.093	.093
	.7	1	.305	.518	.322	.189	.269	.278	.130	.150	.214
	0	.4	.184	.484	.112	.113	.175	.084	.073	.099	.063
	.2	.4	.426	.892	.136	.276	.514	.104	.187	.313	.078
.5	.4	.8	.184	.484	.160	.113	.175	.127	.073	.099	.092
	.7	1	.266	.673	.283	.168	.307	.247	.112	.176	.192
	0	.4	.140	.593	.114	.087	.196	.091	.058	.101	.075
	.2	.4	.328	.811	.116	.213	.535	.092	.146	.354	.073
.75	.4	.8	.140	.593	.140	.087	.196	.111	.058	.101	.086
	.7	1	.203	.733	.226	.129	.401	.188	.088	.193	.152

TABLE 9

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

		n	64			128			256		
ρ	γ	δ	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$
-.5	0	.4	.918	2.88	.164	.480	1.62	.112	.271	.784	.075
	.2	.4	1.78	4.15	.300	.961	1.85	.225	.557	1.03	.159
	.4	.8	.918	2.88	.225	.480	1.62	.161	.271	.784	.108
	.7	1	1.19	3.46	.374	.633	1.70	.296	.363	.886	.216
0	0	.4	1.06	2.98	.192	.553	1.54	.122	.306	.878	.079
	.2	.4	2.04	3.67	.354	1.10	2.12	.253	.634	1.26	.177
	.4	.8	1.06	2.98	.282	.553	1.54	.191	.306	.878	.120
	.7	1	1.37	3.21	.483	.729	1.78	.370	.411	.847	.249
.5	0	.4	.901	3.81	.172	.472	1.55	.115	.266	.877	.075
	.2	.4	1.76	3.59	.319	.953	2.04	.233	.553	1.02	.161
	.4	.8	.901	3.81	.241	.472	1.55	.170	.266	.877	.109
	.7	1	1.17	3.44	.405	.625	1.75	.313	.358	.862	.219
.75	0	.4	.717	2.67	.138	.372	1.30	.093	.212	.914	.066
	.2	.4	1.39	2.94	.248	.747	1.56	.179	.441	.986	.131
	.4	.8	.717	2.67	.195	.372	1.30	.128	.212	.914	.088
	.7	1	.930	2.79	.331	.491	1.50	.232	.286	1.12	.169

TABLE 10

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

		n	64			128			256		
ρ	γ	δ	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$	$\overline{\nu}_I$	$\overline{\nu}_F$	$\overline{\nu}_0$
-.5	0	.4	.608	1.72	.434	.346	.899	.338	.210	.431	.249
	.2	.4	1.09	2.12	.831	.642	1.23	.714	.403	.749	.555
	.4	.8	.608	1.72	1.18	.346	.899	1.04	.210	.431	.818
	.7	1	.761	1.98	2.23	.438	1.03	2.16	.270	.564	1.81
0	0	.4	.666	2.34	.466	.399	.927	.373	.239	.479	.280
	.2	.4	1.14	2.18	.864	.711	1.33	.764	.443	.750	.615
	.4	.8	.666	2.34	1.36	.399	.927	1.18	.239	.479	.907
	.7	1	.813	1.86	2.70	.496	1.10	2.60	.301	.555	2.09
.5	0	.4	.615	1.84	.450	.358	.864	.353	.205	.384	.262
	.2	.4	1.09	2.00	.849	.657	1.25	.729	.408	.691	.585
	.4	.8	.615	1.84	1.24	.358	.864	1.08	.205	.384	.816
	.7	1	.768	1.74	2.38	.451	1.03	2.27	.268	.504	1.85
.75	0	.4	.529	1.70	.383	.295	.774	.297	.166	.325	.217
	.2	.4	.986	1.95	.769	.590	1.27	.652	.362	.651	.508
	.4	.8	.529	1.70	.974	.295	.774	.835	.166	.325	.678
	.7	1	.681	1.72	1.89	.391	.966	1.72	.228	.445	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 $\overline{\nu}_0$
-.5	0	.4	1.78	2.41	.109	1.07	1.53	.084	.670	1.13	.065
	.2	.4	3.46	2.92	.131	2.14	1.79	.102	1.36	1.13	.077
	.4	.8	1.78	2.41	.154	1.07	1.53	.122	.670	1.13	.092
	.7	1	2.30	2.65	.270	1.41	1.50	.237	.887	.999	.188
0	0	.4	2.04	3.09	.107	1.19	1.75	.073	.748	1.23	.049
	.2	.4	4.03	3.57	.141	2.37	2.20	.105	1.52	1.53	.076
	.4	.8	2.04	3.09	.171	1.19	1.75	.128	.748	1.23	.093
	.7	1	2.66	3.51	.322	1.56	1.89	.278	.988	1.25	.214
.5	0	.4	1.74	2.73	.112	1.06	1.82	.084	.668	1.35	.063
	.2	.4	3.39	3.13	.136	2.12	1.90	.104	1.35	1.45	.078
	.4	.8	1.74	2.73	.160	1.06	1.82	.127	.668	1.35	.092
	.7	1	2.26	3.10	.283	1.40	1.73	.247	.881	1.38	.192
.75	0	.4	1.42	2.05	.114	.831	1.60	.091	.519	1.33	.075
	.2	.4	2.74	2.32	.116	1.67	1.47	.092	1.05	1.22	.073
	.4	.8	1.42	2.05	.140	.831	1.60	.111	.519	1.33	.086
	.7	1	1.83	2.17	.226	1.09	1.45	.188	.686	1.20	.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
-.5	0	.4	.076	.137	.072	.091	.068	.059	.124	.174	.124	.142	.122	.102
	.2	.4	.076	.130	.059	.099	.058	.075	.134	.164	.117	.141	.130	.115
	.7	1	.073	.137	.066	.102	.060	.077	.129	.175	.118	.139	.128	.116
0	0	.4	.078	.093	.053	.080	.057	.055	.136	.152	.112	.122	.125	.117
	.2	.4	.077	.055	.054	.033	.062	.034	.133	.082	.104	.073	.114	.066
	.7	1	.076	.074	.058	.060	.053	.055	.134	.128	.105	.102	.120	.099
.5	0	.4	.074	.131	.055	.080	.055	.066	.136	.164	.119	.122	.117	.097
	.2	.4	.073	.114	.055	.094	.054	.081	.141	.146	.120	.135	.111	.108
	.7	1	.068	.115	.055	.083	.050	.080	.140	.154	.121	.127	.116	.116
.75	0	.4	.075	.124	.059	.076	.063	.037	.136	.153	.112	.104	.116	.067
	.2	.4	.073	.170	.058	.146	.069	.093	.143	.207	.113	.183	.116	.137
	.7	1	.076	.145	.060	.111	.064	.075	.143	.178	.113	.148	.110	.116

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	.114	.020	.092	.027	.084	.037	.184	.035	.161	.060	.132	.069
-.5	.2	.4	.109	.029	.098	.037	.074	.049	.180	.053	.158	.066	.138	.082
	.7	1	.112	.029	.097	.032	.082	.044	.182	.047	.161	.065	.136	.078
	0	.4	.122	.024	.080	.022	.077	.028	.187	.053	.150	.044	.129	.066
0	.2	.4	.125	.025	.092	.017	.063	.009	.191	.036	.146	.030	.130	.025
	.7	1	.125	.037	.079	.022	.075	.016	.192	.051	.146	.046	.122	.044
	0	.4	.112	.024	.097	.030	.067	.033	.177	.049	.160	.055	.145	.052
.5	.2	.4	.118	.020	.094	.031	.071	.055	.182	.053	.161	.060	.139	.076
	.7	1	.121	.025	.090	.033	.073	.046	.179	.046	.165	.059	.133	.072
	0	.4	.115	.018	.100	.023	.079	.022	.185	.041	.161	.041	.151	.048
.75	.2	.4	.107	.038	.096	.066	.081	.107	.188	.074	.162	.098	.146	.149
	.7	1	.112	.034	.101	.049	.079	.053	.181	.066	.159	.078	.141	.092

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	.067	.076	.067	.049	.055	.057	.125	.103	.117	.081	.100	.088
-.5	.2	.4	.067	.051	.063	.036	.055	.047	.119	.068	.119	.056	.094	.072
	.7	1	.067	.063	.066	.044	.058	.052	.122	.082	.120	.067	.103	.073
	0	.4	.069	.041	.067	.046	.059	.038	.113	.065	.122	.072	.106	.078
0	.2	.4	.066	.022	.064	.022	.065	.017	.114	.030	.120	.035	.112	.035
	.7	1	.070	.035	.067	.035	.065	.025	.114	.048	.125	.056	.107	.056
	0	.4	.062	.070	.054	.056	.049	.053	.124	.098	.115	.078	.105	.075
.5	.2	.4	.061	.046	.053	.039	.049	.041	.127	.065	.110	.057	.103	.060
	.7	1	.066	.062	.051	.051	.047	.044	.127	.087	.118	.063	.102	.067
	0	.4	.073	.092	.055	.059	.054	.063	.145	.116	.107	.082	.096	.083
.75	.2	.4	.069	.079	.054	.066	.057	.064	.131	.102	.104	.091	.099	.087
	.7	1	.067	.088	.058	.055	.051	.049	.137	.113	.106	.078	.103	.068

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	.101	.036	.081	.030	.068	.030	.171	.059	.140	.055	.115	.070
-.5	.2	.4	.105	.030	.087	.025	.060	.017	.178	.055	.139	.040	.123	.031
	.7	1	.101	.033	.086	.031	.061	.033	.175	.065	.140	.055	.119	.062
	0	.4	.097	.032	.086	.037	.071	.042	.162	.062	.157	.080	.125	.074
0	.2	.4	.090	.031	.091	.025	.077	.020	.166	.057	.150	.044	.127	.039
	.7	1	.092	.034	.089	.034	.070	.030	.155	.056	.150	.066	.124	.056
	0	.4	.112	.037	.073	.028	.053	.038	.165	.054	.141	.057	.101	.071
.5	.2	.4	.097	.019	.078	.029	.064	.018	.161	.044	.139	.051	.120	.045
	.7	1	.109	.026	.082	.030	.060	.034	.164	.050	.147	.062	.110	.059
	0	.4	.117	.025	.082	.022	.051	.026	.185	.047	.133	.048	.104	.060
.75	.2	.4	.107	.022	.078	.026	.065	.024	.173	.044	.133	.038	.114	.037
	.7	1	.111	.019	.081	.029	.058	.031	.184	.047	.143	.058	.106	.053

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	.274	.072	.250	.244	.255	.813	.349	.179	.333	.498	.341	.957
-.5	.2	.4	.258	.001	.228	.009	.228	.085	.331	.017	.317	.053	.317	.239
	.7	1	.270	.025	.243	.068	.233	.392	.343	.056	.331	.215	.334	.682
	0	.4	.258	.477	.245	.807	.248	.988	.344	.621	.319	.882	.325	.995
0	.2	.4	.242	.160	.214	.277	.229	.459	.327	.258	.296	.404	.310	.679
	.7	1	.255	.302	.229	.565	.241	.842	.339	.435	.308	.701	.322	.956
	0	.4	.264	.702	.246	.904	.248	.988	.356	.767	.324	.938	.324	.992
.5	.2	.4	.245	.295	.230	.399	.224	.631	.341	.371	.303	.467	.317	.733
	.7	1	.253	.498	.239	.726	.239	.944	.347	.598	.306	.778	.325	.962
	0	.4	.274	.767	.244	.941	.251	.997	.360	.820	.329	.965	.333	.997
.75	.2	.4	.249	.320	.221	.407	.218	.661	.336	.403	.310	.495	.313	.768
	.7	1	.262	.544	.240	.734	.238	.963	.350	.623	.318	.815	.318	.978

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	.091	.013	.072	.034	.066	.031	.143	.035	.109	.065	.112	.067
-.5	.2	.4	.084	.014	.065	.046	.053	.074	.139	.034	.115	.083	.099	.124
	.7	1	.088	.018	.067	.047	.058	.051	.137	.039	.112	.084	.105	.097
	0	.4	.078	.027	.061	.022	.050	.036	.127	.046	.115	.041	.100	.064
0	.2	.4	.072	.021	.054	.040	.047	.054	.135	.048	.107	.068	.086	.102
	.7	1	.075	.022	.052	.029	.049	.050	.132	.040	.104	.064	.094	.079
	0	.4	.068	.027	.063	.026	.056	.028	.124	.047	.118	.051	.105	.061
.5	.2	.4	.071	.028	.064	.032	.061	.046	.113	.042	.116	.057	.110	.087
	.7	1	.065	.024	.056	.035	.060	.039	.120	.046	.110	.061	.110	.074
	0	.4	.085	.031	.072	.025	.060	.018	.144	.051	.129	.054	.113	.042
.75	.2	.4	.074	.026	.073	.045	.057	.048	.138	.052	.126	.076	.114	.087
	.7	1	.080	.026	.080	.039	.058	.033	.143	.060	.125	.065	.112	.063

TABLE 18

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

n	64			128			256		
estimation$\backslash^{b_2}$	0	.5	.9	0	.5	.9	0	.5	.9
$S_0(d)$	-.016	.403	.852	-.001	.415	.868	-.003	.417	.875
$S_1(d)$	-.325	-.170	.161	-.286	-.121	.114	-.207	-.072	.059

TABLE 19

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

n	64			128			256		
estimation$\backslash^{b_2}$	0	.5	.9	0	.5	.9	0	.5	.9
$S_0(d)$	.110	.149	.107	.072	.102	.075	.046	.070	.057
$S_1(d)$	.419	.258	.286	.409	.216	.237	.373	.182	.165

TABLE 20

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

n	64			128			256		
estimation$\backslash^{b_2}$	0	.5	.9	0	.5	.9	0	.5	.9
$S_0(d)$	.087	.806	1.00	.061	.941	1.00	.085	1.00	1.00
$S_1(d)$	.367	.170	.158	.382	.149	.159	.276	.135	.106

TABLE 21

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

n	64			128			256		
estimation$\backslash^{b_2}$	0	.5	.9	0	.5	.9	0	.5	.9
$S_0(d)$	.130	.837	1.00	.117	.955	1.00	.104	1.00	1.00
$S_1(d)$	.444	.201	.205	.419	.190	.195	.304	.162	.122

TABLE 22

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

estimation	γ	n / $\delta \backslash b$	64: 0	64: .5	64: .9	128: 0	128: .5	128: .9	256: 0	256: .5	256: .9
$S_0(d)$, $T_0(c)$	0	.4	-.038	.396	.842	-.012	.416	.868	-.003	.421	.874
	.2	.4	-.048	.368	.825	-.018	.400	.857	-.005	.413	.870
	.7	1	-.040	.387	.839	-.012	.410	.865	-.005	.418	.872
$S_1(d)$, $T_1(c)$	0	.4	-.250	-.309	.044	-.192	-.218	.066	-.105	-.153	.042
	.2	.4	-.422	-.345	.001	-.361	-.255	.025	-.256	-.177	.012
	.7	1	-.336	-.325	.026	-.279	-.233	.049	-.176	-.164	.030

TABLE 23

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

estimation	γ	n / $\delta \backslash b$	64: 0	64: .5	64: .9	128: 0	128: .5	128: .9	256: 0	256: .5	256: .9
$S_0(d)$, $T_0(c)$	0	.4	.123	.126	.128	.081	.087	.087	.053	.062	.062
	.2	.4	.112	.117	.114	.077	.086	.083	.055	.063	.060
	.7	1	.118	.122	.123	.079	.087	.086	.058	.065	.064
$S_1(d)$, $T_1(c)$	0	.4	.330	.260	.252	.309	.223	.183	.279	.195	.136
	.2	.4	.378	.239	.225	.370	.214	.157	.347	.195	.113
	.7	1	.355	.248	.240	.340	.220	.170	.316	.195	.126

TABLE 24

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

estimation	γ	n / $\delta \backslash b$	64: 0	64: .5	64: .9	128: 0	128: .5	128: .9	256: 0	256: .5	256: .9
$S_0(d)$, $T_0(c)$	0	.4	.114	.951	1.00	.096	1.00	1.00	.075	1.00	1.00
	.2	.4	.105	.946	1.00	.091	1.00	1.00	.084	1.00	1.00
	.7	1	.110	.945	1.00	.075	.999	1.00	.092	1.00	1.00
$S_1(d)$, $T_1(c)$	0	.4	.243	.273	.115	.352	.263	.094	.322	.249	.080
	.2	.4	.388	.301	.108	.471	.297	.074	.455	.271	.068
	.7	1	.329	.274	.104	.400	.268	.089	.362	.254	.073

TABLE 25

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

estimation	γ	n / $\delta \backslash b$	64: 0	64: .5	64: .9	128: 0	128: .5	128: .9	256: 0	256: .5	256: .9
$S_0(d)$, $T_0(c)$	0	.4	.182	.970	1.00	.136	1.00	1.00	.148	1.00	1.00
	.2	.4	.174	.963	1.00	.144	1.00	1.00	.129	1.00	1.00
	.7	1	.163	.958	1.00	.139	1.00	1.00	.182	1.00	1.00
$S_1(d)$, $T_1(c)$	0	.4	.330	.329	.152	.421	.324	.126	.375	.304	.110
	.2	.4	.450	.349	.136	.533	.357	.104	.501	.331	.099
	.7	1	.380	.329	.140	.453	.328	.118	.416	.312	.099

TABLE 26

EFFICIENCY RATIOS

		b			0.5						0.9			
		n	64	128	256	64	128	256	64	128	256	64	128	256
ρ	γ	δ	r_1	r_1	r_1	r_2	r_2	r_2	r_1	r_1	r_1	r_2	r_2	r_2
	0	.4	1.18	1.12	1.15	1.06	1.12	1.10	7.46	5.12	3.63	1.38	3.40	3.14
-.5	.2	.4	5.97	5.00	4.23	1.00	1.43	2.37	45.9	33.6	25.8	1.09	16.3	19.8
	.7	1	1.86	1.54	1.40	.988	1.09	1.15	16.6	11.6	8.52	1.35	7.47	7.30
	0	.4	1.22	1.13	1.12	1.00	1.07	1.05	6.77	5.14	3.60	1.67	3.20	3.26
0	.2	.4	6.11	5.06	4.27	1.07	1.29	2.32	37.9	31.3	23.8	1.91	11.5	20.3
	.7	1	1.97	1.57	1.38	.931	1.02	1.13	14.3	11.3	8.16	2.12	6.82	7.41
	0	.4	1.13	1.14	1.16	1.01	1.11	1.13	7.64	5.10	3.63	1.44	3.69	3.03
.5	.2	.4	6.19	4.61	4.19	1.04	1.35	2.12	45.7	32.8	27.4	.900	4.71	18.8
	.7	1	1.83	1.50	1.39	.961	1.07	1.12	16.9	11.5	8.76	1.26	7.01	7.11
	0	.4	1.16	1.18	1.23	1.16	1.21	1.23	9.14	5.77	3.91	1.53	3.67	2.72
.75	.2	.4	6.11	4.49	4.15	.985	1.39	1.66	61.1	44.2	35.3	.669	15.1	18.5
	.7	1	1.79	1.42	1.38	1.09	1.07	1.01	21.6	14.4	10.4	1.21	7.99	6.80

TABLE 27

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

TABLE 28

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

TABLE 29

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

TABLE 30

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