DEPARTMENT OF ECONOMICS

Working Paper

Local Whittle Analysis of Stationary
Fractional Cointegration

Morten Ørregaard Nielsen

Working Paper No. 2002-8

ISSN 1396-2426

UNIVERSITY OF AARHUS • DENMARK

INSTITUT FOR ØKONOMI

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

WORKING PAPER

Local Whittle Analysis of Stationary
Fractional Cointegration

Morten Ørregaard Nielsen

Working Paper No. 2002-8

DEPARTMENT OF ECONOMICS

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

Local Whittle Analysis of Stationary Fractional Cointegration

(Job Market Paper)

Morten Ørregaard Nielsen*

University of Aarhus

November 12, 2002

Abstract

We consider a local Whittle analysis of a stationary fractionally cointegrated model. A two step estimator, equivalent to the local Whittle QMLE, is proposed to jointly estimate the integration orders of the regressors, the integration order of the errors, and the cointegration vector. The estimator is semiparametric in the sense that it employs local assumptions on the joint spectral density matrix of the regressors and the errors near the zero frequency. We show that, for the entire stationary region of the integration orders, the estimator is asymptotically normal with block diagonal covariance matrix. Thus, the estimates of the integration orders are asymptotically independent of the estimate of the cointegration vector. Furthermore, our estimator of the cointegrating vector is asymptotically normal for a wider range of integration orders than the narrow band frequency domain least squares estimator, and is superior with respect to asymptotic variance. An application to financial volatility series is offered.

JEL Classification: C14; C22

Keywords: Fractional Cointegration; Fractional Integration; Whittle Likelihood; Long Memory; Realized Volatility; Semiparametric Estimation

*Address correspondence to: Morten Ørregaard Nielsen, Department of Economics, Building 322, University of Aarhus, DK-8000 Aarhus C, Denmark; telephone: +45 8942 1584; fax: +45 8613 6334; e-mail: monielsen@econ.au.dk.

1 Introduction

In this paper we are concerned with the joint estimation of the integration orders and the cointegrating vector in stationary fractionally cointegrated models. Suppose we observe the p-vector $z_t = (y_t, x_t')'$, which is integrated of order $d \in (0, 1/2)$, denoted $z_t \in I(d)$. For a precise statement, $z_t \in I(d)$ if

$$(1-L)^d z_t = \varepsilon_t, \tag{1}$$

where $\varepsilon_t \in I(0)$ and $(1-L)^d$ is defined by its binomial expansion

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} L^j, \quad \Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt, \tag{2}$$

in the lag operator L ($Lz_t = z_{t-1}$). A process is labelled $I(0)$ if it is covariance stationary and has spectral density that is bounded and bounded away from zero at the origin.

A scalar-valued stochastic process generated by (1) has spectral density

$$f(\lambda) \sim g\lambda^{-2d} \quad \text{as } \lambda \to 0^+, \tag{3}$$

where g is a constant and the symbol "$\sim$" means that the ratio of the left- and right-hand sides tends to one in the limit. Such a process is said to possess strong dependence or long range dependence, since the autocorrelations decay at a hyperbolic rate in contrast to the much faster exponential rate in the weak dependence case. The parameter d determines the memory of the process. If $d > -1/2$, z_t is invertible and admits a linear representation, and if $d < 1/2$ it is covariance stationary. If $d = 1/2$, the spectral density is bounded at the origin, and the process has only weak dependence. Sometimes, z_t is said to have intermediate memory, short memory, and long memory when $d < 0$, $d = 0$, and $d > 0$, respectively.

Suppose further that $z_t = (y_t, x_t')'$ satisfies the regression model

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

where the error term is integrated of a smaller order $d_e < d$, i.e. $e_t \in I(d_e)$. A much studied special case is the standard $I(1) - I(0)$ cointegration model which arises when $d = 1$ and

$d_e = 0$, see e.g. Watson (1994) for a review. When d and/or d_e are not integers the model is called a fractional cointegration model following the original idea by Granger (1981). We call the model (4) with $0 \leq d_e < d < 1/2$ a stationary fractionally cointegrated model, since it is concerned with the long-run linear co-movement between two or more stationary fractionally integrated processes. The properties of the model in the standard $I(1) - I(0)$ cointegration case are well known, see Watson (1994), but the fractional cointegration framework has been examined only recently, see the short review in Robinson and Yajima (2002).

Since our model is stationary, a comparison with the standard time series regression model with weakly dependent regressors is natural. The new complication is that, since the regressors and the errors both have long memory, they are potentially correlated even at very long horizons, thus rendering the OLS estimator inconsistent, see Robinson (1994) and Robinson and Marinucci (1998). To deal with this issue, Robinson (1994) proposed a semiparametric narrow band frequency domain least squares (FDLS) estimator that assumes only a multivariate generalization of (3), and essentially performs OLS on a degenerating band of frequencies around the origin. The consistency of the estimator in the stationary case is proved by Robinson (1994), and Christensen and Nielsen (2001) show that its asymptotic distribution is normal when the collective memory of the regressors and the error term is less than 1/2, i.e. when $d + d_e < 1/2$. In contrast, Robinson and Marinucci (1998) consider several cases where the regressors are fractionally integrated and nonstationary, and show that the limiting distributions for the FDLS estimator are then functionals of fractional Brownian motion.

Throughout this paper, we shall be concerned with the case $d \in (0, 1/2)$. This interval is relevant for many applications in finance, e.g. stock market trading volume (Lobato and Velasco (2000)), exchange rate volatility (Andersen, Bollerslev, Diebold and Labys (2001)), stock return volatility (Andersen, Bollerslev, Diebold and Ebens (2001) and Christensen and Nielsen (2001)), and spot prices for crude oil (Robinson and Yajima (2002)). In particular, it is the relevant region for the volatility processes we study below in our empirical application.

Many estimators of the memory parameter d and the scale parameter g have been suggested in the literature. A semiparametric approach has been developed by Geweke and Porter-Hudak (1983), Robinson (1994, 1995a, 1995b), Lobato and Robinson (1996), and Lobato (1999), among others. The semiparametric estimators of the memory parameter assume only the model (3) for the spectral density, and use a degenerating part of the periodogram around the origin to estimate the model. This approach has the advantage of being invariant to any short- and medium-term dynamics (as well as mean terms since the zero frequency is usually left out). In particular, a local Whittle QMLE approach based on the maximization of a local Whittle approximation to the likelihood, see our equation (7), has been developed by Robinson (1995a) (who called it a Gaussian semiparametric estimator) and Lobato (1999) to estimate the integration orders of univariate and multivariate stationary fractionally integrated time series, respectively. Of course, a fully parametric approach is more efficient, using the entire sample, but is inconsistent if the parametric model is specified incorrectly, e.g. if the lag-structure of the short-term dynamics is misspecified.

The methods described above are combined by Marinucci and Robinson (2001) and Christensen and Nielsen (2001), who suggest conducting a fractional cointegration analysis in several steps. First, the integration orders of the raw data is estimated by, e.g., the local Whittle QMLE. Secondly, the narrow band FDLS estimator for the cointegrating vector is calculated, and finally the integration order of the residuals is estimated assuming that the approach is equally valid for residuals. Hypothesis testing is then conducted on d_e as if e_t were observed, and on β as if d_e (which enters in the limiting distribution of the FDLS estimator) were known. Although this may indeed be a valid course of action, see Hassler, Marmol and Velasco (2000) and Velasco (2001), a joint estimation method for the integration orders and the cointegration vector would be preferable.

We propose a simple joint semiparametric two step estimator of the integration orders and the cointegration vector in (4), which is equivalent to the local Whittle QMLE. Similarly to the narrow band FDLS estimator for the cointegration vector and the local Whittle QMLE of the integration orders, our estimator employs local assumptions on the joint spectral density matrix of the regressors and the errors near the zero frequency. It turns out that the limiting distribution of our estimator has a block diagonal covariance matrix, so that the estimates of the integration orders are asymptotically uncorrelated with the estimates of the cointegration vector. Thus, the limiting distribution of the estimates of the integration orders equals that derived by Lobato (1999), and in particular, it is unaffected by the fact that it is based in part on residuals. In contrast to the FDLS estimator, we show that our estimator is asymptotically normal for the entire parameter space, i.e. $0 \leq d_e < d < 1/2$, thus avoiding the condition $d + d_e < 1/2$ required by the FDLS estimator for asymptotic normality. We also demonstrate that our estimator, in addition to being applicable for a wider range of integration orders, has smaller asymptotic variance than the FDLS estimator when the latter is asymptotically normal. A similar approach to ours is considered by Velasco (2001) for bivariate nonstationary fractionally cointegrated processes, and similar results for the asymptotic distribution are reached using data tapering, following Lobato and Velasco (2000). However, his results are limited to a bivariate model, and require tapering and an additional user chosen bandwidth parameter to trim out the very first Fourier frequencies as in Robinson (1995*b*).

Following the semiparametric approach outlined above, our estimator enjoys the extremely general treatment of the short-term dynamics that has made the log-periodogram and local Whittle estimators popular among practitioners. In particular, the short-term dynamics does not even need to be specified, since only a degenerating band of frequencies around the origin is used. In contrast, for a parametric estimator to be consistent we would have to specify correctly the short-run dynamics of the model, employing e.g. a vector fractional ARIMA specification as in Dueker and Startz (1998). The obvious cost for this robustness is that the efficiency of the semiparametric estimator relative to a correctly specified parametric estimator converges to zero.

The stationary fractional cointegration model has many potential applications, especially in finance. Many financial time series, like the volatility of stock returns and exchange rates, have been found to be well described by stationary fractionally integrated processes, see e.g.

Andersen, Bollerslev, Diebold and Ebens (2001), Andersen, Bollerslev, Diebold and Labys (2001), and Christensen and Nielsen (2001). Our model then applies if it is assumed that this is a common trend between two or more such processes, which would seem like a plausible assumption especially if the underlying assets are traded on the same market (exchange rate or stock market).

To illustrate our new procedure, we offer an application to the relation between the volatility implied by option prices and the volatility subsequently realized in the stock market. The unbiasedness hypothesis in the option market implies a slope coefficient of unity in the implied-realized volatility relation, but the ordinary regression estimate is less than one-half. However, we conduct a stationary fractional cointegration analysis, and find that the volatility series are well described as being stationary fractionally cointegrated with d approximately 0.45 and d_e insignificantly different from zero. When accounting for the possibility of stationary fractional cointegration, the estimated slope coefficient is insignificantly different from unity, thus supporting long-run unbiasedness of implied volatility as a forecaster of realized volatility.

The paper is organized as follows. In the next section we present the model and set up the local Whittle likelihood and the assumptions necessary to prove our main result. In section 3 we state our main result on the asymptotic distribution of the joint semiparametric two step estimator, and compare this to the local Whittle QMLE of d and the narrow band FDLS estimator of β. Section 4 presents the empirical application to the implied-realized volatility relation, and section 5 concludes. The proof of the main theorem is provided in two appendices.

2 Stationary Fractional Cointegration Model

Let us now generalize the simple model described above. In particular, suppose the spectral density matrix of the p-vector $w_t = (x_t', e_t)'$ is

$$f(\lambda) = \Lambda^{-1} G \Lambda^{-1} \quad \text{as } \lambda \to 0^+, \tag{5}$$

where $\Lambda = \text{diag}(\lambda^{d_1}, ..., \lambda^{d_p})$, $d_a \in \Delta = \{x| \, 0 \leq x \leq \Delta_1, 0 < \Delta_1 < 1/2\}$, $a = 1, ..., p$, and G is a $p \times p$ real symmetric matrix. Equation (5) is the natural multivariate extension of (3), including multivariate fractional ARIMA models as a special case, and is also considered in previous work by e.g. Robinson (1995*b*), Lobato (1999), and Robinson and Yajima (2002). Thus, the elements of the vector x_t can be integrated of different orders, i.e. $x_{at} \in I(d_a)$. This implies, by (4), that $y_t \in I(\max_{1 \leq a \leq p-1} d_a)$, such that the conceptual requirement that at least two of the variables in $(y_t, x_t')'$ must be integrated of the same order is automatically satisfied. Notice that d_p is now the integration order of the error term, i.e. $e_t \in I(d_p)$.

We collect the parameters of interest in the $(2p-1)$-vector $\theta = \left(d_1, ..., d_p, \beta'\right)'$. The Whittle approximation to the (negative) likelihood is (see Lobato (1999))

$$W(\theta, G) = \int_{-\pi}^{\pi} \left(\log |f(\lambda)| + \text{tr}\left[f^{-1}(\lambda) \, \text{Re}\,(I(\lambda))\right]\right) d\lambda,$$

where $I(\lambda) = (2\pi n)^{-1} \left|\sum_{t=1}^{n} w_t e^{it\lambda}\right|^2$ is the periodogram matrix of w_t at frequency λ. In the spirit of the semiparametric approach, we prefer the discrete local version of the likelihood

$$\bar{W}(\theta, G) = \frac{1}{m} \sum_{j=1}^{m} \left(\log |f(\lambda_j)| + \text{tr}\left[f^{-1}(\lambda_j) \, \text{Re}\,(I(\lambda_j))\right]\right) \tag{6}$$

evaluated at the Fourier frequencies $\lambda_j = 2\pi j/n, j = 1, ..., m$. We let the bandwidth parameter $m = m(n)$ tend to infinity to gather information, but at a slower rate than n to remain in a neighborhood of $\lambda = 0$. Note that the zero frequency has been left out of the summation in (6) to render the estimation invariant to mean terms. An integral version of (6) could also have been considered, but it would not share this property and it would be computationally more burdensome.

The local Whittle estimator of (θ, G) is defined as

$$(\hat{\theta}, \hat{G}) = \arg\min_{\theta, G} \bar{W}(\theta, G)$$

over a compact subset of $\Delta^p \times \mathbb{R}^{p^2+p-1}$. We concentrate G out of the likelihood by setting $\hat{G}(\theta) = m^{-1} \sum_{j=1}^{m} \Lambda_j \, \text{Re}(I(\lambda_j)) \Lambda_j$, and write the concentrated likelihood as

$$L(\theta) = \log \left|\hat{G}(\theta)\right| - \frac{2\left(\sum_{a=1}^{p} d_a\right)}{m} \sum_{j=1}^{m} \log \lambda_j \tag{7}$$

apart from constants. The local Whittle estimator of the parameter of interest, θ, can then be defined in terms of the concentrated likelihood as

$$\hat{\theta} = \arg\min_{\theta\in\Theta} L(\theta), \tag{8}$$

where the minimization is carried out over Θ, a compact subset of $\Delta^p \times \mathbb{R}^{p-1}$.

We propose the following simple two step estimator (TSE) for the integration orders and the cointegrating vector,

$$\hat{\theta}^{(2)} = \hat{\theta}^{(1)} - \left(\left.\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}\right|_{\hat{\theta}^{(1)}} \right)^{-1} \left(\left.\frac{\partial L(\theta)}{\partial\theta}\right|_{\hat{\theta}^{(1)}} \right), \tag{9}$$

where $\hat{\theta}^{(1)}$ is a consistent initial estimator, e.g. the local Whittle QMLE of Robinson (1995*a*) and Lobato (1999) combined with the narrow band FDLS estimator of Robinson (1994), Robinson and Marinucci (1998), and Christensen and Nielsen (2001). We could iterate (9) until convergence for higher order gains, but that does not change the first order asymptotics. It is well known that the TSE has the same asymptotic distribution as the QMLE, but we prefer the TSE for its simplicity.

To prove our main result we assume, with obvious implications for y_t, the following conditions on $w_t = (x_t', e_t)'$, the bandwidth, and the initial estimates.

Assumption 1 *The spectral density matrix of w_t given in* (5) *with typical element $f_{ab}(\lambda)$, the cross spectral density between w_{at} and w_{bt}, satisfies*

$$\left| f_{ab}(\lambda) - g_{ab}\lambda^{-d_a-d_b} \right| = O\left(\lambda^{\alpha-d_a-d_b}\right) \text{ as } \lambda \to 0^+, \ a, b = 1, ..., p,$$

for some $\alpha \in (0, 2]$. The matrix G satisfies $g_{ap} = g_{pa} = 0$ for $a = 1, ..., p-1$, and the leading $(p-1)\times(p-1)$ submatrix of G, denoted $\bar{G}$, is positive definite.

Assumption 2 *w_t is a linear process, $w_t = \mu + \sum_{j=0}^{\infty} A_j\varepsilon_{t-j}$, with square summable coefficient matrices, $\sum_{j=0}^{\infty} \|A_j\|^2 < \infty$. The innovations satisfy, almost surely, $E(\varepsilon_t | \mathcal{F}_{t-1}) = 0$, $E(\varepsilon_t\varepsilon_t' | \mathcal{F}_{t-1}) = I_p$, and the matrices $\mu_3 = E(\varepsilon_t \otimes \varepsilon_t\varepsilon_t' | \mathcal{F}_{t-1})$ and $\mu_4 = E(\varepsilon_t\varepsilon_t' \otimes \varepsilon_t\varepsilon_t' | \mathcal{F}_{t-1})$ are nonstochastic, finite, and do not depend on t, where $\mathcal{F}_t = \sigma(\{\varepsilon_s, s \le t\})$.*

Assumption 3 *As* $\lambda \to 0^+$,

$$\frac{dA_a(\lambda)}{d\lambda} = O\left(\lambda^{-1} \|A_a(\lambda)\|\right),\ a = 1, ..., p,$$

where $A_a(\lambda)$ *is the* a'*th row of* $A(\lambda) = \sum_{j=0}^{\infty} A_j e^{ij\lambda}$.

Assumption 4 *The bandwidth parameter* $m = m(n)$ *satisfies*

$$\frac{1}{m} + \frac{m^{1+2\alpha} (\log m)^2}{n^{2\alpha}} \to 0 \text{ as } n \to \infty.$$

Assumption 5 *The initial estimates* $\hat{\theta}^{(1)}$ *are consistent, and in particular satisfy*

$$\begin{aligned} \hat{d}_a^{(1)} - d_a &= O_p\left(m^{-1/2}\right) \text{ for } a = 1, ..., p, \\ \hat{\beta}_a^{(1)} - \beta_a &= O_p\left(m^{-1/2} \lambda_m^{d_a - d_p}\right) \text{ for } a = 1, ..., p-1. \end{aligned}$$

The first part of Assumption 1 specializes (5) by imposing smoothness conditions on the spectral density matrix of w_t commonly employed in the literature. They are satisfied with $\alpha = 2$ if, for instance, w_t is a vector fractional ARIMA process. The condition that $\bar{G}$ be positive definite is a no multicollinearity or no cointegration condition within the components of x_t. The condition that $g_{ap} = g_{pa} = 0$, for $a = 1, ..., p-1$, is new compared to previous research, and ensures that the coherence between the regressors and the error process is of smaller order at the origin. The condition can be thought of as a local-to-zero version of the usual orthogonality condition from least squares theory, and is needed, for instance, to show that the estimation of d_p is unaffected by the fact that it is based in part on estimated residuals.

Assumptions 2 and 3 follow Robinson (1995*a*) and Lobato (1999) in imposing a linear structure on w_t with square summable coefficients and martingale difference innovations with finite fourth moments. Assumption 2 is satisfied, for instance, if ε_t is an *i.i.d.* process with finite fourth moments. Under Assumption 2 we can write the spectral density matrix of w_t as

$$f(\lambda) = \frac{1}{2\pi} A(\lambda) A^*(\lambda), \tag{10}$$

where the asterisk is complex conjugation combined with transposition.

Assumption 4 restricts the expansion rate of the bandwidth parameter $m = m(n)$. The bandwidth is required to tend to infinity for consistency, but at a slower rate than n to remain in a neighborhood of the origin, where we have some knowledge of the form of the spectral density. When α is high, (5) is a better approximation to (10) as $\lambda \to 0^+$, and hence (by the second term of Assumption 4) a higher expansion rate of the bandwidth can be chosen. The weakest constraint is implied by $\alpha = 2$, in which case the condition is $m = o(n^{4/5})$.

Finally, Assumption 5 states the required rates of convergence of the initial estimates. It is satisfied, for instance, if the integration orders are estimated by the local Whittle QMLE or the log-periodogram method (see Robinson (1995*a*, 1995*b*) and Lobato (1999) for the estimation of d for observed data, and Hassler et al. (2000) and Velasco (2001) for estimation of d for residuals) and the cointegration vector by the FDLS estimator (see Christensen and Nielsen (2001)), but other estimators would also satisfy this assumption.

3 Main Result

We are now ready to state our main result.

Theorem 1 *Let θ_0 denote the true value of the parameter vector θ, and suppose θ_0 belongs to the interior of the parameter space, Θ. Under $0 \leq d_p < d_a < 1/2$, for $a = 1, ..., p-1$, (4), and Assumptions 1-5*

$$\sqrt{m}\,\mathrm{diag}\left(I_p, \lambda_m^{d_p} \bar{\Lambda}_m^{-1}\right)\left(\hat{\theta}^{(2)} - \theta_0\right) \xrightarrow{D} N\left(0, \Omega^{-1}\right), \tag{11}$$

with

$$\Omega = \begin{pmatrix} E & 0 \\ 0 & F \end{pmatrix}, \tag{12}$$

$$E = 2\left(I_p + G \odot G^{-1}\right), \tag{13}$$

$$F_{ab} = \frac{2g_{ab}}{g_{pp}\left(1 - d_a - d_b + 2d_p\right)} \quad a, b = 1, ..., p-1, \tag{14}$$

where $\odot$ *denotes the Hadamard product and* $\bar{\Lambda}_m$ *is the leading* $(p-1)\times(p-1)$ *submatrix of* $\Lambda_m = \mathrm{diag}(\lambda_m^{d_a}, ..., \lambda_m^{d_p})$.

Proof. The asymptotic distribution of the TSE is the same as that of the QMLE, which is given by (11) if we can show the following. The score is such that

$$\sqrt{m}\,\mathrm{diag}\left(I_p, \lambda_m^{-d_p}\bar{\Lambda}_m\right)\frac{\partial L\left(\theta_0\right)}{\partial\theta} \xrightarrow{D} N\left(0,\Omega\right), \tag{15}$$

and the Hessian satisfies

$$\mathrm{diag}\left(I_p, \lambda_m^{-d_p}\bar{\Lambda}_m\right)\frac{\partial^2 L(\tilde{\theta})}{\partial\theta\partial\theta'}\,\mathrm{diag}\left(I_p, \lambda_m^{-d_p}\bar{\Lambda}_m\right) \xrightarrow{p} \Omega \tag{16}$$

for all $\tilde{\theta}$ such that $\parallel \tilde{\theta}-\theta_0 \parallel \leq \parallel \hat{\theta}^{(1)}-\theta_0 \parallel$. Notice that Ω is positive definite by Assumption 1 and the fact that the Hadamard product of two positive definite matrices is positive definite. We prove (15) in appendix A, where parts of this proof follow Lobato (1999) in applying the martingale difference array approximation technique by Robinson (1995*a*). (16) is proven in appendix B. ■

Some comments on our result are in order. Velasco (2001) reaches a result very similar to our Theorem 1 in his nonstationary setup, using tapered periodograms to account for the nonstationarity, following the approach of Lobato and Velasco (2000). However, his results are limited to a bivariate model, and require tapering and an additional user chosen bandwidth parameter (say l) to trim out the first l Fourier frequencies as in Robinson (1995*b*).

The asymptotic distribution in (11) is block diagonal, such that the estimates of the integration orders are asymptotically uncorrelated with the estimate of the cointegration vector. In particular, the asymptotic distribution of the estimators of the integration orders is unaffected by the fact that they are based in part on residuals. This is due to the local orthogonality condition in Assumption 1, which ensures that the effect of the estimation of β on the estimation of the integration orders is negligible. A discussion of the efficiency gains of the multivariate estimator of the integration orders over the univariate local Whittle QMLEs in Robinson (1995*a*) can be found in Lobato (1999, p. 136).

Let us have a closer look at the asymptotic distribution in the simple two variable case. Suppose we observe two time series y_t and x_t both integrated of order $d < 1/2$, and that the error term is known to be integrated of order $d_e < d$. Then the asymptotic distribution (11) of $\hat{\beta}$ in Theorem 1 reduces to

$$\sqrt{m}\lambda_m^{d_e-d}\left(\hat{\beta}^{(2)} - \beta_0\right) \xrightarrow{D} N\left(0, \frac{g_e\left(1-2d+2d_e\right)}{2g_x}\right),$$

where g_x and g_e are the elements of G, which is a diagonal 2×2 matrix. Thus, the variance depends on the signal-to-noise ratio g_x/g_e.

We compare our estimator of the cointegration vector with the narrow band FDLS given by

$$\hat{\beta}_{FDLS} = \left(\frac{1}{m}\sum_{j=1}^{m}\operatorname{Re}\left(I_{xx}\left(\lambda_j\right)\right)\right)^{-1}\frac{1}{m}\sum_{j=1}^{m}\operatorname{Re}\left(I_{xy}\left(\lambda_j\right)\right), \tag{17}$$

and asymptotically distributed according to

$$\sqrt{m}\lambda_m^{d_e-d}\left(\hat{\beta}_{FDLS} - \beta_0\right) \xrightarrow{D} N\left(0, \frac{g_e\left(1-2d\right)^2}{2g_x\left(1-2d-2d_e\right)}\right)$$

in the two variable case with $d + d_e < 1/2$, see Christensen and Nielsen (2001). We note immediately that the convergence rates are the same, and in particular, they are very close to $\sqrt{n}$ for relevant parameter values. For instance, when $m = O\left(n^{0.6}\right)$ and $d - d_e = 0.4$, which are values close to those in the empirical application below, we get that $\hat{\beta}^{(2)}$ is $n^{0.46}$-consistent. The asymptotic relative efficiency of $\hat{\beta}^{(2)}$ with respect to $\hat{\beta}_{FDLS}$ is

$$\frac{V(\hat{\beta}_{FDLS})}{V(\hat{\beta})} = \frac{(1-2d)^2}{(1-2d)^2 - 4d_e^2},$$

which equals unity if and only if $d_e = 0$, and exceeds unity otherwise. Thus, our estimator is more efficient and applies for a wider range of (d, d_e) than the FDLS estimator.

The unknown parameters appearing in the asymptotic distribution (11) can be replaced by consistent estimates. In particular, the matrix of coherencies at the zero frequency, G, can be estimated by $\hat{G}(\hat{\theta}^{(2)})$, which is consistent by Lobato (1999, p. 136). Its asymptotic distribution could also be derived by application of the delta-method, see Robinson (1995a) and Lobato and Velasco (2000).

Based on Theorem 1 it is straightforward to construct Wald tests of hypotheses that involve both the integration orders and the cointegration vector. For instance, the linear restrictions $H_0 : R\theta = r$ can be tested by

$$W = \left(R\hat{\theta} - r\right)' \left(R\Omega^{-1}R'\right)^{-1} \left(R\hat{\theta} - r\right) \xrightarrow{D} \chi_q^2 \tag{18}$$

under the null, where q is the number of linearly independent restrictions. Some hypotheses of general interest in this framework are (i) the components of x_t are integrated of the same order, $\theta_1 = ... = \theta_{p-1}$, (ii) the errors have no long memory, $\theta_p = 0$, (iii) x_{kt} is not present in the cointegrating relation, $\theta_{p+k} = 0$, or combinations of these.

4 Empirical Application

In this section we conduct an actual stationary fractional cointegration analysis of the relation between the volatility implied by option prices, and the subsequent realized return volatility of the underlying asset, following Christensen and Nielsen (2001).

If option market participants are rational and markets are efficient, the price of a financial option should reflect all publicly available information including information about future return volatility of the underlying asset. Given an observation on the price of an option, the implied volatility σ_{IV} may be determined by inverting the option pricing formula with respect to σ_{IV}, and if this is done every period t a time series $\sigma_{IV,t}$ results. Each implied volatility $\sigma_{IV,t}$ may now be considered as the market's forecast of the actually realized return volatility of the underlying asset. Here, realized volatility is simply the sample standard deviation $\sigma_{RV,t}$ of the realized return from t to $t + 1$. In practice, we work with the log volatilities, since they are close to Gaussian, see Andersen, Bollerslev, Diebold and Ebens (2001).

Christensen and Prabhala (1998) considered the regression specification

$$y_t = \alpha + \beta x_t + e_t, \tag{19}$$

where $y_t = \ln \sigma_{RV,t}$ and $x_t = \ln \sigma_{IV,t}$ are the log volatilities, and α and β are intercept and slope coefficients. The unbiasedness hypothesis for option markets implies a β-coefficient of unity. A monthly sampling frequency was employed for x_t and y_t. The underlying asset was the S&P100 stock market index, and y_t was calculated from daily returns, see Christensen and Prabhala (1998) for the details. Basic OLS regression in (19) produced a β-estimate that was significantly greater than zero and less than unity (Christensen and Prabhala (1998) also presented results without the log transform and the difference was negligible).

Inferences from OLS may be erroneous if x_t and y_t are fractionally cointegrated, which is exactly what would be expected under the unbiasedness hypothesis. Thus, if $x_t{=}E_t(y_t)$ with $E_t(\cdot)$ denoting conditional expectation as of time t, then β is unity and e_t is serially uncorrelated. For a detailed description of the implied-realized volatility relation and its implications, see Christensen and Prabhala (1998). If volatility is fractionally integrated, as empirical literature suggests (Andersen, Bollerslev, Diebold and Ebens (2001) and Christensen and Nielsen (2001) find fractional integration with d around $0.35-0.45$), whereas the forecasting error e_t in (19) possesses only short memory, then x_t and y_t are fractionally cointegrated. This is in fact what the empirical results in Christensen and Nielsen (2001) and our empirical results below indicate. In particular, Christensen and Nielsen (2001) considered a fractional cointegration analysis of (19), using first the univariate local Whittle estimator of Robinson (1995a) to estimate the integration orders of the raw data, then the narrow band FDLS estimator to estimate β, and finally the local Whittle estimator to estimate the integration order of the errors. It was found that accounting for the possibility of stationary fractional cointegration greatly improves the results, and in most cases produces β-estimates that are insignificantly different from unity.

The data we use are the same as those investigated by Christensen and Nielsen (2001), and are weekly data covering the period January 1, 1988, to December 31, 1995, resulting in $n = 417$ observations. The final data series are based on high-frequency data from the Berkeley Options Data Base (BODB), see the BODB User's Guide for a description. From the high-frequency options data, a 5-minute return series for the underlying S&P500 index is constructed for the period 9:00 AM to 3:00 PM each trading day. This results in a series of 147,022 observations. From this 5-minute return series we form the realized volatility $\sigma_{RV,t}$ over each one-week interval by taking the sample standard deviation of the 5-minute annualized returns in week t.

The implied volatilities are backed out from the Monday 10:00 AM quote, for the call of shortest maturity and closest to the money, using the standard option pricing formula corrected for dividends. This results in a weekly implied volatility series $\tilde{\sigma}_{IV,t}$ with different times to maturity since the options expire monthly. We convert this heterogeneous series to another weekly series $\sigma_{IV,t}$, that may be associated with the series $\sigma_{RV,t}$ of realized volatilities covering homogeneous nonoverlapping weekly intervals by the formula

$$\sigma^2_{IV,t-i} = \frac{1}{d_i - d_{i-1}} \left(d_i \cdot \tilde{\sigma}^2_{IV,t-i} - d_{i-1} \cdot \tilde{\sigma}^2_{IV,t-i+1} \right), \tag{20}$$

where d_i is the number of days until expiration of $\tilde{\sigma}_{IV,t-i}$, starting with $\sigma_{IV,t} = \tilde{\sigma}_{IV,t}$ for t corresponding to a one-week option and then applying the recursion (20). This is of course an approximation for implied volatilities, as opposed to realized volatilities where it is an identity. However, the approximation is a high-frequency measurement error, and consequently our semiparametric approach should be robust towards it. For the complete details of the construction of the data set and summary statistics, see Christensen and Nielsen (2001).

In Tables 1-3 we report the results of our stationary fractional cointegration analysis for bandwidths $m = n^{0.5} = 20$, $m = n^{0.55} = 27$, and $m = n^{0.6} = 37$, respectively.

The first column in each table shows the initial estimates. For the integration orders d and d_e we choose Robinson's (1995a) univariate local Whittle estimates using the same bandwidth parameter as for the TSE. To estimate β we choose the FDLS estimator and, following Robinson and Marinucci (1998) and Marinucci and Robinson (2001), we use a lower bandwidth for the FDLS estimator, and in particular the bandwidth $m = 5$ was used. The results are robust to changes in this bandwidth parameter, at least up to about 15, see also Christensen and Nielsen (2001). The initial estimates are comparable to those found by Christensen and Nielsen (2001), and in particular the series seem to be stationary ($d < 1/2$) and the errors are close to $I(0)$.

The initial estimate of β is 0.866, which is well above the $0.3 - 0.4$ that are typical for OLS estimates of β, see Christensen and Prabhala (1998) and Christensen and Nielsen (2001), but still suggests that implied volatility is a biased forecast of realized volatility.

Tables 1-3 about here

In the next two columns we report the TSE (9) and the standard error of each parameter, respectively. The standard errors are calculated using our new distribution theory as the square root of the diagonal elements of the covariance matrix in Theorem 1. For d the estimates are virtually unchanged compared to the initial estimates of approximately 0.45, which is in line with previous evidence, see Andersen, Bollerslev, Diebold and Ebens (2001) and Christensen and Nielsen (2001). Turning to the estimation of the parameters of primary interest, d_e and β, we find much smaller point estimates of d_e, ranging from 0.04 to 0.07 compared to the initial estimates of 0.08 to 0.10, and much larger estimates of β ranging from 1.12 to 1.19 depending on the bandwidth, but the estimates of β are insignificantly different from unity.

The results so far are consistent with the notion that realized and implied volatility are well described as stationary but fractionally integrated series, and that they tend to move together in the sense that the errors in (19) have less memory. The interesting question is how closely they move together and whether the errors are in fact only weakly dependent. To answer this question, the fourth column shows the Wald test statistic (18) of the joint hypothesis that $d_e = 0$ and $\beta = 1$, which is asymptotically distributed as a χ^2 random variable with 2 degrees of freedom (the 5% and 1% critical values are 5.99 and 9.21, respectively). The test does not reject for any choice of bandwidth, suggesting that implied and realized volatility can indeed be described by a stationary fractionally cointegrated relation with unit coefficient and only weakly dependent errors. Thus, the results indicate that all long memory properties in volatility are common features for implied and realized volatility.

The remaining columns in Tables 1-3 presents the estimates, standard errors, and Wald test statistic when (9) is iterated until convergence to five decimal places. These results do not differ greatly from the results for the TSE. The estimates of d are virtually unchanged compared to the TSE. For the integration order of the errors, the estimates are less than 0.0001 for all bandwidths, obviously strongly in favor of the hypothesis of weakly dependent errors. The estimates of β are closer to unity now, ranging from 1.01 to 1.05 and leaving the possibility of a unit coefficient highly likely. Consequently, the Wald statistics for the converged estimates are insignificant for all choices of bandwidth.

Thus, similarly to Andersen, Bollerslev, Diebold and Ebens (2001) and Christensen and Nielsen (2001), we find that the volatility series are well described as stationary fractionally integrated series. From the residual analysis we cannot reject that implied and realized volatility indeed are stationary fractionally cointegrated. That is, the residuals are of lower order of fractional integration than the volatility series themselves, $d_e < d$. In fact, our results are consistent with the even stronger relation that $d_e = 0$. Under long-run unbiasedness, we would expect the series to follow each other closely resulting in a unit β-coefficient, which is also supported by our analysis. Hence, the relation between implied and realized volatility indeed appears to be one of stationary fractional cointegration.

5 Conclusion

We consider a local Whittle analysis of a stationary fractionally cointegrated model. In particular, we propose a two step estimator, which is equivalent to the local Whittle QMLE, to jointly estimate the integration orders of the regressors, the integration order of the errors, and the cointegration vector. The estimator is semiparametric in the sense that it employs local assumptions on the joint spectral density matrix of the regressors and the errors near the zero frequency, following the approach by Robinson (1995a) and Lobato (1999) for estimating the integration orders. By using a degenerating part of the periodogram near the origin, the approach is invariant to short-run dynamics, which would have to be specified correctly in a parametric procedure.

In our stationary fractionally integrated case, we show that the two step estimator is asymptotically normal with block diagonal covariance matrix for the entire stationary region of the integration orders. Thus, the estimates of the integration orders are asymptotically uncorrelated with the estimate of the cointegration vector. Furthermore, our estimator of the cointegration vector is asymptotically normal for a wider range of integration orders than the narrow band frequency domain least squares estimator of Robinson (1994), analyzed by Robinson and Marinucci (1998), Marinucci and Robinson (2001), and Christensen and Nielsen (2001), and is superior with respect to asymptotic variance when the latter is normal.

To demonstrate the feasibility of our methodology in practice, we have offered an application to financial volatility series. The unbiasedness hypothesis of option markets implies a coefficient of unity in the implied-realized volatility relation, but the ordinary regression estimate is less than one-half. We show that implied and realized volatility are well described as being stationary fractionally cointegrated. When accounting for this, our estimates of this coefficient are more than twice as large as before and insignificantly different from unity. Furthermore, we are unable to reject the joint hypothesis of weak dependence of the error process and unit coefficient in any of our specifications. This demonstrates that useful long-run relations can be derived even among stationary series.

Acknowledgements

I would like to thank Richard Blundell, Bent Jesper Christensen, Niels Haldrup, Uwe Hassler, Helmut Lütkepohl, Herman van Dijk, and participants at the 2002 Econometric Society European Meeting in Venice and the 2002 Econometric Society Winter European Meeting in Budapest for many helpful comments and discussions. The first version of the paper was completed while I was visiting Yale University and the Cowles Foundation, their hospitality is gratefully acknowledged.

Appendix A: Limit of the Score

Applying the Cramér-Wold device we need to show that

$$\eta' \sqrt{m}\, \mathrm{diag}\left(I_p, \lambda_m^{-d_p} \bar{\Lambda}_m\right) \frac{\partial L(\theta_0)}{\partial \theta} \xrightarrow{D} N\left(0, \eta' \Omega \eta\right) \tag{21}$$

for any non-null vector η. The derivatives with respect to d_a and β_a are

$$\frac{\partial L(\theta_0)}{\partial d_a} = \frac{2}{m} \sum_{j=1}^{m} \nu_j \,\mathrm{Re}\left(g^a \Lambda_j I_{wa}(\lambda_j) \lambda_j^{d_a} - 1\right), \tag{22}$$

$$\frac{\partial L(\theta_0)}{\partial \beta_a} = -\frac{2}{m} \sum_{j=1}^{m} \lambda_j^{d_p} \,\mathrm{Re}\left(g^p \Lambda_j I_{wa}(\lambda_j)\right), \tag{23}$$

where (22) is the same as in Lobato (1999), $\nu_j = \log j - m^{-1} \sum_{j=1}^{m} \log j$, g^a is the a'th row of G^{-1} and $I_{wa}(\lambda)$ is the cross-periodogram between w_t and w_{at}. In both (22) and (23) we replaced $\hat{G}(\theta_0)$ by G since

$$\left\| \hat{G}(\theta_0) - G \right\| = O_p\left(m^{-1/2}\right), \tag{24}$$

see Lobato (1999).

The part of the left-hand side of (21) corresponding to (23) is

$$\begin{aligned}
& -\sum_{a=1}^{p-1} \eta_{a+p} \lambda_m^{d_a - d_p} \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \lambda_j^{d_p} \,\mathrm{Re}\left(g^p \Lambda_j I_{wa}(\lambda_j)\right) \\
= & -\sum_{a=1}^{p-1} \eta_{a+p} \lambda_m^{d_a - d_p} \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \lambda_j^{d_p} \,\mathrm{Re}\left(g^p \Lambda_j \left(I_{wa}(\lambda_j) - A(\lambda_j) J(\lambda_j) A_a^*(\lambda_j)\right)\right) \qquad (25) \\
& -\sum_{a=1}^{p-1} \eta_{a+p} \lambda_m^{d_a - d_p} \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \lambda_j^{d_p} \,\mathrm{Re}\left(g^p \Lambda_j A(\lambda_j) J(\lambda_j) A_a^*(\lambda_j)\right), \qquad (26)
\end{aligned}$$

where $J(\lambda)$ is the periodogram of ε_t and $A_a(\lambda)$ is the a'th row of $A(\lambda)$. By (C.2) of Lobato (1999), which is implied by our assumptions,

$$\begin{aligned}
(25) &= O_p\left(\sum_{a=1}^{p-1} \frac{1}{\sqrt{m}} \left(m^{1/3} (\log m)^{2/3} + \log m + \frac{\sqrt{m}}{n^{1/4}}\right)\right) \\
&= O_p\left(\frac{(\log m)^{2/3}}{m^{1/6}} + \frac{\log m}{\sqrt{m}} + \frac{1}{n^{1/4}}\right) \xrightarrow{p} 0.
\end{aligned}$$

Write (26) as

$$
\begin{aligned}
&-\sum_{a=1}^{p-1}\eta_{a+p}\lambda_m^{d_a-d_p}\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\lambda_j^{d_p}\operatorname{Re}\left(g^p\Lambda_j A(\lambda_j)\frac{1}{2\pi n}\left|\sum_{t=1}^{n}\varepsilon_t e^{it\lambda_j}\right|^2 A_a^*(\lambda_j)\right)\\
= &-\sum_{a=1}^{p-1}\eta_{a+p}\lambda_m^{d_a-d_p}\frac{1}{\pi\sqrt{m}}\sum_{j=1}^{m}\lambda_j^{d_p}\operatorname{Re}\left(g^p\Lambda_j A(\lambda_j)A_a^*(\lambda_j)\right) \qquad (27)\\
&-\sum_{a=1}^{p-1}\eta_{a+p}\lambda_m^{d_a-d_p}\frac{1}{\pi\sqrt{m}}\sum_{j=1}^{m}\lambda_j^{d_p}\operatorname{Re}\left(g^p\Lambda_j A(\lambda_j)\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t\varepsilon_t'-I_p\right)A_a^*(\lambda_j)\right) \qquad (28)\\
&-\sum_{a=1}^{p-1}\eta_{a+p}\lambda_m^{d_a-d_p}\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\lambda_j^{d_p}\operatorname{Re}\left(g^p\Lambda_j A(\lambda_j)\frac{1}{2\pi n}\sum_{t=1}^{n}\sum_{s\neq t}\varepsilon_t\varepsilon_s' e^{i(t-s)\lambda_j}A_a^*(\lambda_j)\right). \qquad (29)
\end{aligned}
$$

By definition of $f(\lambda)$, see (10), and using Assumption 1

$$
\begin{aligned}
(27) &= \max_{1\le a\le p-1} O\left(\frac{1}{\sqrt{m}}\lambda_m^{d_a-d_p}\sum_{j=1}^{m}\lambda_j^{2d_p}f_{pa}(\lambda_j)\right)\\
&= \max_{1\le a\le p-1} O\left(\frac{1}{\sqrt{m}}\lambda_m^{d_a-d_p}\sum_{j=1}^{m}\lambda_j^{\alpha-d_a+d_p}\right),
\end{aligned}
$$

which is $O\left(n^{-2\alpha}m^{1+2\alpha}\right)\to 0$ by Assumption 4. For equation (28), note that $\varepsilon_t\varepsilon_t'-I_p$ is a martingale difference sequence with respect to $\mathcal{F}_t$ implying that $n^{-1}\sum_{t=1}^{n}\varepsilon_t\varepsilon_t'-I_p=O_p(n^{-1/2})$. Thus,

$$
\begin{aligned}
(28) &= \max_{1\le a\le p-1} O_p\left(\frac{1}{\sqrt{m}}\lambda_m^{d_a-d_p}\sum_{j=1}^{m}\lambda_j^{2d_p}\frac{1}{\sqrt{n}}f_{pa}(\lambda_j)\right)\\
&= \max_{1\le a\le p-1} O_p\left(\frac{1}{\sqrt{nm}}\lambda_m^{d_a-d_p}\sum_{j=1}^{m}\lambda_j^{\alpha-d_a+d_p}\right)\\
&= O_p\left(\lambda_m^{1/2+\alpha}\right).
\end{aligned}
$$

We are left with (29), which we rewrite as

$$
\sum_{t=1}^{n}\varepsilon_t'\sum_{s=1}^{t-1}\sum_{a=1}^{p-1}\frac{\eta_{a+p}}{\pi n\sqrt{m}}\lambda_m^{d_a-d_p}\sum_{j=1}^{m}\lambda_j^{d_p}\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^{p\prime}e^{i(t-s)\lambda_j}\bar{A}_a(\lambda_j)\right)\varepsilon_s.
$$

The corresponding term for (22), derived by Lobato (1999, p. 141), is given by

$$
\sum_{t=1}^{n}\varepsilon_t'\sum_{s=1}^{t-1}\sum_{a=1}^{p-1}\frac{\eta_a}{\pi n\sqrt{m}}\sum_{j=1}^{m}\lambda_j^{d_a}\nu_j\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^{a\prime}e^{i(t-s)\lambda_j}\bar{A}_a(\lambda_j)\right)\varepsilon_s.
$$

Thus, (21) has the same asymptotic distribution as $\sum_{t=1}^{n}\varepsilon_t'\sum_{s=1}^{t-1}c_{t-s,n}\varepsilon_s$, where we define

$$\begin{aligned}
c_{tn} &= \frac{1}{\pi n\sqrt{m}}\sum_{j=1}^{m}\left(\theta_{j1}+\theta_{j2}\right)\cos\left(t\lambda_j\right),\\
\theta_{j1} &= \nu_j\sum_{a=1}^{p}\lambda_j^{d_a}\eta_a\operatorname{Re}\left(A'\left(\lambda_j\right)\Lambda_j g^{a\prime}\bar{A}_a\left(\lambda_j\right)+A_a'\left(\lambda_j\right)g^a\Lambda_j\bar{A}\left(\lambda_j\right)\right),\\
\theta_{j2} &= -\lambda_j^{d_p}\sum_{a=1}^{p-1}\eta_{a+p}\lambda_m^{d_a-d_p}\operatorname{Re}\left(A'\left(\lambda_j\right)\Lambda_j g^{p\prime}\bar{A}_a\left(\lambda_j\right)+A_a'\left(\lambda_j\right)g^p\Lambda_j\bar{A}\left(\lambda_j\right)\right).
\end{aligned}$$

Notice that, by construction, $\|\theta_{j1}\|=O(1)$ and $\|\theta_{j2}\|=O\left(\max_{1\le a\le p-1}\left(m/j\right)^{d_a-d_p}\right)$.

Since $z_{tn}=\varepsilon_t'\sum_{s=1}^{t-1}c_{t-s,n}\varepsilon_s$ is a martingale difference array with respect to $\mathcal{F}_t=\sigma\left(\{\varepsilon_s,s\le t\}\right)$, we can apply the CLT for martingale difference arrays if (see Hall and Heyde (1980, chp. 3.2))

$$\sum_{t=1}^{n}E\left(z_{tn}^2\middle|\mathcal{F}_{t-1}\right)-\sum_{a=1}^{2p-1}\sum_{b=1}^{2p-1}\eta_a\eta_b\Omega_{ab}\xrightarrow{p}0, \tag{30}$$

$$\sum_{t=1}^{n}E\left(z_{tn}^2 1\left(|z_{tn}|>\delta\right)\right)\to 0 \text{ for all } \delta>0. \tag{31}$$

A sufficient condition for (31) is

$$\sum_{t=1}^{n}E\left(z_{tn}^4\right)\to 0. \tag{32}$$

First, to show (30),

$$\begin{aligned}
\sum_{t=1}^{n}E\left(z_{tn}^2\middle|\mathcal{F}_{t-1}\right) &= \sum_{t=1}^{n}E\left(\sum_{s=1}^{t-1}\sum_{r=1}^{t-1}\varepsilon_s'c_{t-s,n}'\varepsilon_t\varepsilon_t'c_{t-r,n}\varepsilon_r\middle|\mathcal{F}_{t-1}\right)\\
&= \sum_{t=1}^{n}\sum_{s=1}^{t-1}\varepsilon_s'c_{t-s,n}'c_{t-s,n}\varepsilon_s \qquad (33)\\
&\quad+\sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{r\ne s}\varepsilon_s'c_{t-s,n}'c_{t-r,n}\varepsilon_r. \qquad (34)
\end{aligned}$$

The term (34) has mean zero and variance

$$O\left(n\left(\sum_{s=1}^{n}\|c_{sn}\|^2\right)^2+\sum_{t=3}^{n}\sum_{u=2}^{t-1}\left(\sum_{s=1}^{u-1}\|c_{u-s,n}\|^2\sum_{s=1}^{u-1}\|c_{t-s,n}\|^2\right)\right) \tag{35}$$

by (D.10) and (D.11) of Lobato (1999). It is immediate that $\|c_{sn}\|=O\left(\left(n\sqrt{m}\right)^{-1}\sum_{j=1}^{m}\|\theta_{j1}+\theta_{j2}\|\right)=O\left(n^{-1}m^{1/2}\log m\right)$, using that $\sup_{-1\le\alpha\le C}\left|m^{-\alpha-1}\left(\log m\right)^{-1}\sum_{j=1}^{m}j^{\alpha}\right|=O(1)$ for $C\in(1,\infty)$.

Another bound is

$$\begin{aligned}
\|c_{sn}\| &= O\left(\frac{1}{n\sqrt{m}}\sum_{j=1}^{m}\|\theta_{j1}+\theta_{j2}\|\left|\cos\left(s\lambda_j\right)\right|\right) \\
&= O\left(\frac{m^{d_a-d_p-1/2}}{n}\sum_{j=1}^{m} j^{d_p-d_a}\left|\cos\left(s\lambda_j\right)\right|\right) \\
&= O\left(\frac{m^{d_a-d_p-1/2}}{n}\left(m^{1+2d_p-2d_a}\right)^{1/2}\left(\frac{n}{s}\right)^{1/2}\right) \\
&= O\left(\frac{1}{\sqrt{ns}}\right),
\end{aligned}$$

where the third line follows by the Cauchy-Schwartz Inequality and the relation $\sum_{j=1}^{k}\left|\cos\left(s\lambda_j\right)\right|^2 \leq \sum_{j=1}^{k}\left|\cos\left(s\lambda_j\right)\right| = O\left(n/s\right)$. This bound is better when $s > n/m$. Thus,

$$\begin{aligned}
\sum_{s=1}^{n}\|c_{sn}\|^2 &= O\left(\sum_{s=1}^{[n/m]}\frac{m\left(\log m\right)^2}{n^2} + \sum_{s=[n/m]+1}^{n}\frac{1}{ns}\right) \\
&= O\left(\frac{\left(\log m\right)^2}{n} + \frac{\log n}{n}\right),
\end{aligned}$$

implying that the first term of (35) is $O\left(n^{-1}\log^2 n\right)$. The second term of (35) is

$$O\left(n\left(\sum_{s=1}^{n}\|c_{sn}\|^2\right)\left(\sum_{s=1}^{[n/2]} s\,\|c_{sn}\|^2\right)\right),$$

following the analysis in Lobato (1999, p. 151) and Robinson (1995*a*, p. 1646-1647). A third bound for $\|c_{sn}\|$ is

$$\begin{aligned}
\|c_{sn}\| &= O\left(\frac{m^{d_a-d_p-1/2}}{n}\sum_{j=1}^{m}\left|\cos\left(s\lambda_j\right)\right|\right) \\
&= O\left(\frac{m^{d_a-d_p-1/2}}{s}\right),
\end{aligned}$$

using the above relations. Applying this bound, we find that

$$\begin{aligned}
\sum_{s=1}^{[n/2]} s\,\|c_{sn}\|^2 &= O\left(\sum_{s=1}^{[n/2]}\frac{m^{2d_a-2d_p-1}}{s}\right) \\
&= O\left(m^{2d_a-2d_p-1}\log n\right)
\end{aligned}$$

and (35) $= O\left(n^{-1}(\log n)^2 + m^{2d_a-2d_p-1}(\log n)^2\right)$.

We still need to show that the mean of (33) is asymptotically equal to $\sum_{a=1}^{2p-1}\sum_{b=1}^{2p-1}\eta_a\eta_b\Omega_{ab}$. Thus,

$$\begin{aligned} E\,(33) &= \sum_{t=1}^{n}\sum_{s=1}^{t-1} E\,\mathrm{tr}\left(c_{t-s,n}'c_{t-s,n}\varepsilon_s\varepsilon_s'\right) \\ &= \sum_{t=1}^{n}\sum_{s=1}^{t-1} \mathrm{tr}\left(c_{t-s,n}'c_{t-s,n}\right) \end{aligned}$$

by Assumption 2. Rewrite this expression as

$$\sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{j=1}^{m}\sum_{j'=1}^{m}\frac{1}{\pi^2n^2m}\,\mathrm{tr}\left(\left(\theta_{j1}'+\theta_{j2}\right)\left(\theta_{j'1}'+\theta_{j'2}\right)\right)\cos\left(\left(t-s\right)\lambda_j\right)\cos\left(\left(t-s\right)\lambda_{j'}\right)$$

$$= \sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{j=1}^{m}\frac{1}{\pi^2n^2m}\,\mathrm{tr}\left(\theta_{j1}'\theta_{j1}\right)\cos^2\left(\left(t-s\right)\lambda_j\right) \tag{36}$$

$$+\sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{j=1}^{m}\frac{1}{\pi^2n^2m}\,\mathrm{tr}\left(\theta_{j2}'\theta_{j2}\right)\cos^2\left(\left(t-s\right)\lambda_j\right) \tag{37}$$

$$+\sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{j=1}^{m}\frac{1}{\pi^2n^2m}2\,\mathrm{tr}\left(\theta_{j1}'\theta_{j2}\right)\cos^2\left(\left(t-s\right)\lambda_j\right) \tag{38}$$

$$+\sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{j=1}^{m}\sum_{j'\neq j}^{m}\frac{1}{\pi^2n^2m}\,\mathrm{tr}\left(\theta_{j1}'\theta_{j'1}\right)\cos\left(\left(t-s\right)\lambda_j\right)\cos\left(\left(t-s\right)\lambda_{j'}\right) \tag{39}$$

$$+\sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{j=1}^{m}\sum_{j'\neq j}^{m}\frac{1}{\pi^2n^2m}\,\mathrm{tr}\left(\theta_{j2}'\theta_{j'2}\right)\cos\left(\left(t-s\right)\lambda_j\right)\cos\left(\left(t-s\right)\lambda_{j'}\right) \tag{40}$$

$$+\sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{j=1}^{m}\sum_{j'\neq j}^{m}\frac{1}{\pi^2n^2m}2\,\mathrm{tr}\left(\theta_{j1}'\theta_{j'2}\right)\cos\left(\left(t-s\right)\lambda_j\right)\cos\left(\left(t-s\right)\lambda_{j'}\right). \tag{41}$$

It was shown by Lobato (1999) that (36) is asymptotically equal to $\sum_{a=1}^{p}\sum_{b=1}^{p}\eta_a\eta_bE_{ab}$ and that (39) is asymptotically negligible. We consider the remaining terms in turn. First,

$$(40) = \max_{1\le a\le p-1} O\left(\sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{j=1}^{m}\sum_{j'\neq j}^{m}\frac{1}{n^2m}\left(\frac{m}{j}\right)^{d_a-d_p}\left(\frac{m}{j'}\right)^{d_a-d_p}\cos\left(\left(t-s\right)\lambda_j\right)\cos\left(\left(t-s\right)\lambda_{j'}\right)\right)$$

and, using that $\sum_{t=1}^{n}\sum_{s=1}^{t-1}\cos\left(\left(t-s\right)\lambda_j\right)\cos\left(\left(t-s\right)\lambda_{j'}\right) = -n/2$ for $\lambda_j \neq \lambda_{j'}$, we can bound (40) by $\max_{1\le a\le p-1} O\left(\left(nm\right)^{-1}m^{2d_a-2d_p}\sum_{j=1}^{m}j^{d_p-d_a}\sum_{j'\neq j}^{m}j'^{d_p-d_a}\right) = O\left(m/n\right)$. Similarly, (41) is also $O\left(m/n\right)$.

For the covariance term in (38) we notice that

$$
\begin{aligned}
\operatorname{tr}\left(\frac{1}{4\pi^2}\theta_{j1}'\theta_{j2}\right) &= -\nu_j \sum_{a=1}^{p}\sum_{b=1}^{p-1}\eta_a\eta_{b+p}\lambda_m^{d_b-d_p}\lambda_j^{d_a+d_p} \\
&\quad\times\left[\operatorname{tr}\left(\frac{1}{4\pi^2}\operatorname{Re}\left(A_a'(\lambda_j)g^a\Lambda_j\bar{A}(\lambda_j)\right)\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^{p\prime}\bar{A}_b(\lambda_j)\right)\right)\right. \\
&\quad+\operatorname{tr}\left(\frac{1}{4\pi^2}\operatorname{Re}\left(A_a'(\lambda_j)g^a\Lambda_j\bar{A}(\lambda_j)\right)\operatorname{Re}\left(A_b'(\lambda_j)g^p\Lambda_j\bar{A}(\lambda_j)\right)\right) \\
&\quad+\operatorname{tr}\left(\frac{1}{4\pi^2}\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^{a\prime}\bar{A}_a(\lambda_j)\right)\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^{p\prime}\bar{A}_b(\lambda_j)\right)\right) \\
&\quad\left.+\operatorname{tr}\left(\frac{1}{4\pi^2}\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^{a\prime}\bar{A}_a(\lambda_j)\right)\operatorname{Re}\left(A_b'(\lambda_j)g^p\Lambda_j\bar{A}(\lambda_j)\right)\right)\right]
\end{aligned}
$$

and, using the definition of $f(\lambda)$ and Assumption 1, this is easily shown to be $o(1)$. E.g., the first term in square brackets is $\operatorname{tr}\left(f_{ab}(\lambda_j)g^a\Lambda_j f(\lambda_j)\Lambda_j g^{p\prime}\right)=O\left(\lambda_j^{-d_a-d_b}g^a\lambda_j^{\alpha}\right)=O\left(\lambda_j^{\alpha-d_a-d_b}\right)$ using that $g_{ap}=0$ for $a=1,...,p-1$. This implies that (38) is $o(1)$ since $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos^2(s\lambda_j)=(n-1)^2/4$.

Now let us examine $\operatorname{tr}(\theta_{j2}'\theta_{j2})$ appearing in (37),

$$
\begin{aligned}
\operatorname{tr}\left(\frac{\theta_{j2}'\theta_{j2}}{4\pi^2}\right) &= \operatorname{tr}\left(\sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\frac{\lambda_m^{d_a+d_b-2d_p}\lambda_j^{2d_p}}{4\pi^2}\operatorname{Re}\left(A_a'(\lambda_j)g^p\Lambda_j\bar{A}(\lambda_j)\right)\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^{p\prime}\bar{A}_b(\lambda_j)\right)\right) \\
&\quad+\operatorname{tr}\left(\sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\frac{\lambda_m^{d_a+d_b-2d_p}\lambda_j^{2d_p}}{4\pi^2}\operatorname{Re}\left(A_a'(\lambda_j)g^p\Lambda_j\bar{A}(\lambda_j)\right)\operatorname{Re}\left(A_b'(\lambda_j)g^p\Lambda_j\bar{A}(\lambda_j)\right)\right) \\
&\quad+\operatorname{tr}\left(\sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\frac{\lambda_m^{d_a+d_b-2d_p}\lambda_j^{2d_p}}{4\pi^2}\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^p\bar{A}_a(\lambda_j)\right)\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^{p\prime}\bar{A}_b(\lambda_j)\right)\right) \\
&\quad+\operatorname{tr}\left(\sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\frac{\lambda_m^{d_a+d_b-2d_p}\lambda_j^{2d_p}}{4\pi^2}\operatorname{Re}\left(A'(\lambda_j)\Lambda_j g^p\bar{A}_a(\lambda_j)\right)\operatorname{Re}\left(A_b'(\lambda_j)g^p\Lambda_j\bar{A}(\lambda_j)\right)\right).
\end{aligned}
$$

By definition of $f(\lambda)$, the first term is asymptotically equal to $\sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\lambda_j^{2d_p}f_{ba}(\lambda_j)\times g^p\Lambda_j f(\lambda_j)\Lambda_j g^{p\prime}=\sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\lambda_j^{2d_p-d_a-d_b}g_{ba}g_{pp}^{-1}$, the fourth to $\sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\lambda_j^{2d_p-d_a-d_b}g_{ab}g_{pp}^{-1}$, and the second and third terms to zero using Assumption 1. Hence, (37) is asymptotically equal to

$$
\sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{j=1}^{m}\sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\frac{4\lambda_m^{d_a+d_b-2d_p}\lambda_j^{2d_p-d_a-d_b}}{n^2 m}\frac{(g_{ab}+g_{ba})}{g_{pp}}\cos^2((t-s)\lambda_j). \qquad (42)
$$

We can approximate the Riemann sum appearing in (42) by an integral, viz.

$$\frac{2\pi}{n}\sum_{j=1}^{m}\lambda_j^{2d_p-d_a-d_b}\sim\int_0^{\lambda_m}\lambda^{2d_p-d_a-d_b}d\lambda=\frac{\lambda_m^{1-d_a-d_b+2d_p}}{1-d_a-d_b+2d_p},$$

where the symbol "$\sim$" means that the ratio of the left- and right-hand sides tends to one. Using this approximation we get that

$$\begin{aligned}(42) \quad &\sim \sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\left(\sum_{t=1}^{n}\sum_{s=1}^{t-1}\cos^2\left((t-s)\lambda_j\right)\right)\frac{2\lambda_m^{d_a+d_b-2d_p}}{\pi nm}\frac{g_{ab}+g_{ba}}{g_{pp}}\frac{\lambda_m^{1-d_a-d_b+2d_p}}{1-d_a-d_b+2d_p}\\ &= \sum_{a=1}^{p-1}\sum_{b=1}^{p-1}\eta_{a+p}\eta_{b+p}\frac{2g_{ab}}{g_{pp}\left(1-d_a-d_b+2d_p\right)},\end{aligned}$$

since $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos^2\left(s\lambda_j\right)=\left(n-1\right)^2/4$, and we have shown (30).

Thus, we need to show (32),

$$\begin{aligned}\sum_{t=1}^{n}E\left(z_{tn}^4\right) &= \sum_{t=1}^{n}E\left(\sum_{s=1}^{t-1}\varepsilon_s'c_{t-s,n}\varepsilon_t\varepsilon_t'\sum_{r=1}^{t-1}c_{t-r,n}\varepsilon_r\sum_{p=1}^{t-1}\varepsilon_p'c_{t-p,n}\varepsilon_t\varepsilon_t'\sum_{q=1}^{t-1}c_{t-q,n}\varepsilon_q\right)\\ &\leq C\left(\sum_{t=1}^{n}\mathrm{tr}\left(\sum_{s=1}^{t-1}c_{t-s,n}'c_{t-s,n}c_{t-s,n}'c_{t-s,n}\right)+\sum_{t=1}^{n}\mathrm{tr}\left(\sum_{s=1}^{t-1}c_{t-s,n}'\sum_{r=1}^{t-1}c_{t-r,n}c_{t-r,n}'c_{t-s,n}\right)\right)\end{aligned}$$

for some constant $C>0$ by Assumption 2. This expression can be bounded by $O\left(n\left(\sum_{t=1}^{n}\left\|c_{tn}^2\right\|\right)^2\right)=O\left(n^{-1}(\log n)^2\right)$, and we are done.

Appendix B: Limit of the Hessian

We prove that

$$\frac{\partial^2L(\tilde{\theta})}{\partial d_a\partial d_b}\xrightarrow{p}E_{ab}, \tag{43}$$

$$\lambda_m^{d_b-d_p}\frac{\partial^2L(\tilde{\theta})}{\partial d_a\partial\beta_b}\xrightarrow{p}0, \tag{44}$$

$$\lambda_m^{d_a+d_b-2d_p}\frac{\partial^2L(\tilde{\theta})}{\partial\beta_a\partial\beta_b}\xrightarrow{p}F_{ab}, \tag{45}$$

for all $\tilde{\theta}$ such that $\|\tilde{\theta}-\theta_0\|\leq\|\hat{\theta}^{(1)}-\theta_0\|$.

First, we will need to strengthen the approximation (24) to G by showing that

$$\left\|\hat{G}(\tilde{\theta})-\hat{G}\left(\theta_{0}\right)\right\|=O_{p}\left(\frac{\log n}{\sqrt{m}}\right). \tag{46}$$

The proof for the leading $(p-1)\times(p-1)$ block is given in Lobato (1999, pp. 145-148). Consider now, for $a=1,...,p-1$,

$$\hat{g}_{ap}(\tilde{\theta})-\hat{g}_{ap}\left(\theta_{0}\right)=\frac{1}{m}\sum_{j=1}^{m}\left(\lambda_{j}^{\tilde{d}_{a}+\tilde{d}_{p}}\tilde{I}_{ap}\left(\lambda_{j}\right)-\lambda_{j}^{d_{a}+d_{p}}I_{ap}\left(\lambda_{j}\right)\right),$$

where $\tilde{I}_{ap}(\lambda)$ is the cross-periodogram between $\tilde{e}_t=y_t-\tilde{\beta}'x_t$ and x_{at}. Noting that $\tilde{I}_{ap}(\lambda)-I_{ap}(\lambda)=(\beta-\tilde{\beta})'I_{xa}(\lambda)$, we can rewrite this as

$$\hat{g}_{ap}(\tilde{\theta})-\hat{g}_{ap}\left(\theta_{0}\right)=\frac{1}{m}\sum_{j=1}^{m}\lambda_{j}^{\tilde{d}_{a}+\tilde{d}_{p}}(\beta-\tilde{\beta})'I_{xa}\left(\lambda_{j}\right)+\frac{1}{m}\sum_{j=1}^{m}\left(\lambda_{j}^{\tilde{d}_{a}+\tilde{d}_{p}}-\lambda_{j}^{d_{a}+d_{p}}\right)I_{ap}\left(\lambda_{j}\right). \tag{47}$$

The first term on the right-hand side can be bounded as

$$\begin{aligned}\frac{1}{m}\sum_{j=1}^{m}\lambda_{j}^{\tilde{d}_{a}+\tilde{d}_{p}}\sum_{b=1}^{p-1}(\beta_{b}-\tilde{\beta}_{b})I_{ba}\left(\lambda_{j}\right) &\leq \frac{1}{m}\left(\max_{1\leq j\leq m}\lambda_{j}^{\tilde{d}_{a}+\tilde{d}_{p}-d_{a}-d_{p}}\right)\sum_{j=1}^{m}\lambda_{j}^{d_{a}+d_{p}}\sum_{b=1}^{p-1}(\beta_{b}-\tilde{\beta}_{b})I_{ba}\left(\lambda_{j}\right)\\ &= O_{p}\left(m^{-1/2}\right),\end{aligned}$$

using $\max_{1\leq j\leq m}\lambda_{j}^{\tilde{d}_{a}+\tilde{d}_{p}-d_{a}-d_{p}}=O_{p}(1)$ and Assumption 5. The second term on the right-hand side of (47) is

$$O_{p}\left(\frac{1}{m}\left(\max_{1\leq j\leq m}\lambda_{j}^{\tilde{d}_{a}+\tilde{d}_{p}-d_{a}-d_{p}}-1\right)\sum_{j=1}^{m}\lambda_{j}^{d_{a}+d_{p}}I_{ap}\left(\lambda_{j}\right)\right)=O_{p}\left(\frac{\log n}{\sqrt{m}}\right)$$

by Assumption 5 and the above analysis. The $(p,p)'$th element of (46) follows in the exact same way by application of the Cauchy-Schwartz Inequality.

In view of (46), (43) follows from Lobato (1999). For (44) and (45) it can be shown that

$$\lambda_{m}^{d_{b}-d_{p}}\left(\frac{\partial^{2}L(\tilde{\theta})}{\partial d_{a}\partial\beta_{b}}-\frac{\partial^{2}L\left(\theta_{0}\right)}{\partial d_{a}\partial\beta_{b}}\right)\xrightarrow{p}0, \tag{48}$$

$$\lambda_{m}^{d_{a}+d_{b}-2d_{p}}\left(\frac{\partial^{2}L(\tilde{\theta})}{\partial\beta_{a}\partial\beta_{b}}-\frac{\partial^{2}L\left(\theta_{0}\right)}{\partial\beta_{a}\partial\beta_{b}}\right)\xrightarrow{p}0, \tag{49}$$

by proceeding component by component with the same methods that we applied to show (46).

We show next that

$$\lambda_m^{d_a+d_b-2d_p} \frac{\partial^2 L(\theta_0)}{\partial \beta_a \partial \beta_b} \xrightarrow{p} F_{ab}. \tag{50}$$

The left-hand side of (50) is asymptotically equal to

$$\lambda_m^{d_a+d_b-2d_p} \frac{2}{m} \sum_{j=1}^m \lambda_j^{d_p} \operatorname{Re}\left(g^p \Lambda_j \begin{pmatrix} 0_{p-1} \\ I_{ab}(\lambda_j) \end{pmatrix} \right) \tag{51}$$

$$-\lambda_m^{d_a+d_b-2d_p} \frac{2}{m} \sum_{j=1}^m \lambda_j^{d_p} \operatorname{Re}\left(g^p \left(\frac{1}{m} \sum_{j'=1}^m \Lambda_{j'} \begin{pmatrix} O_{p-1} & I_{xa}(\lambda_{j'}) \\ I_{ax}(\lambda_{j'}) & 2I_{pa}(\lambda_{j'}) \end{pmatrix} \right) \Lambda_{j'} G^{-1} \Lambda_j I_{wa}(\lambda_j) \right) \tag{52}$$

by (24), with 0_{p-1} and O_{p-1} denoting a $(p-1)$-vector of zeros and a $(p-1)\times(p-1)$ matrix of zeros, respectively. The first of these terms is

$$\begin{aligned}(51) \quad = \quad & \lambda_m^{d_a+d_b-2d_p} \frac{2}{m} \sum_{j=1}^m \lambda_j^{2d_p} \operatorname{Re}\left(g^{pp}\left(I_{ab}(\lambda_j) - A_a(\lambda_j) J(\lambda_j) A_b^*(\lambda_j)\right)\right) \\ & + \lambda_m^{d_a+d_b-2d_p} \frac{2}{m} \sum_{j=1}^m \lambda_j^{2d_p} \operatorname{Re}\left(g^{pp} A_a(\lambda_j) J(\lambda_j) A_b^*(\lambda_j)\right),\end{aligned}$$

where the first term is $o_p(1)$ by the same arguments as for (25) in appendix A. The second term is

$$\begin{aligned} & \lambda_m^{d_a+d_b-2d_p} \frac{2}{m} \sum_{j=1}^m \lambda_j^{2d_p} \operatorname{Re}\left(g^{pp} A_a(\lambda_j) \frac{1}{2\pi n} \left| \sum_{t=1}^n \varepsilon_t e^{it\lambda_j} \right|^2 A_b^*(\lambda_j) \right) \\ = \quad & \lambda_m^{d_a+d_b-2d_p} \frac{2}{m} \sum_{j=1}^m \lambda_j^{2d_p} \frac{1}{2\pi} \operatorname{Re}\left(g^{pp} A_a(\lambda_j) A_b^*(\lambda_j)\right) + o_p(1) \end{aligned} \tag{53}$$

by Assumption 2 and the same arguments as for (27) – (28) in appendix A.

By definition of $f(\lambda)$ we get that

$$(53) = \lambda_m^{d_a+d_b-2d_p} \frac{2}{m} \sum_{j=1}^m \lambda_j^{2d_p} g^{pp} \operatorname{Re}\left(f_{ab}(\lambda_j)\right) + o_p(1)$$

and, applying the integral approximation from appendix A, this expression is asymptotically equal to

$$\lambda_m^{d_a+d_b-2d_p} \frac{n}{m\pi} g^{pp} \int_0^{\lambda_m} \lambda^{2d_p} \left(g_{ab} \lambda^{-d_a-d_b} \right) d\lambda = \frac{2 g_{ab}}{g_{pp}\left(1 - d_a - d_b + 2d_p\right)}.$$

Next, rewrite (52) as

$$-\lambda_m^{d_a+d_b-2d_p}\frac{2}{m}\sum_{j=1}^{m}\lambda_j^{d_p}\operatorname{Re}\left(\frac{1}{m}\sum_{j'=1}^{m}\left(\begin{array}{cc} g^{pp}\lambda_{j'}^{d_p}I_{ax}\left(\lambda_{j'}\right), & g^{p}\Lambda_{j'}I_{wa}\left(\lambda_{j'}\right)+g^{pp}\lambda_{j'}^{d_p}I_{pa}\left(\lambda_{j'}\right)\end{array}\right)\Lambda_{j'}G^{-1}\Lambda_j I_{wa}\left(\lambda_j\right)\right)$$

$$=O_p\left(\lambda_m^{d_a+d_b-2d_p}\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{d_p}\left(\frac{1}{m}\sum_{j'=1}^{m}\lambda_{j'}^{d_p+d_a}\right)\lambda_j^{-d_a}\right)$$

applying the same type of analysis as in appendix A. The last expression is seen to be $O(\lambda_m^{d_a+d_b})=o\left(1\right)$.

To complete the proof, we need to show that

$$\lambda_m^{d_b-d_p}\frac{\partial^2 L\left(\theta_0\right)}{\partial d_a\partial\beta_b}\xrightarrow{p}0, \qquad (54)$$

which implies (44) in view of (48). The left-hand side of (54) is asymptotically equal to

$$\lambda_m^{d_b-d_p}\frac{2}{m}\sum_{j=1}^{m}\nu_j\operatorname{Re}\left(g^a\left(\frac{1}{m}\sum_{j'=1}^{m}\Lambda_{j'}\left(\begin{array}{cc} O_{p-1} & I_{xa}\left(\lambda_{j'}\right)\\ I_{ax}\left(\lambda_{j'}\right) & 2I_{pa}\left(\lambda_{j'}\right)\end{array}\right)\right)\Lambda_{j'}G^{-1}\Lambda_j I_{wa}\left(\lambda_j\right)\lambda_j^{d_a}\right)$$

$$-\lambda_m^{d_b-d_p}\frac{2}{m}\sum_{j=1}^{m}\nu_j\operatorname{Re}\left(g^a\Lambda_j\left(\begin{array}{c}0_{p-1}\\ I_{ab}\left(\lambda_j\right)\end{array}\right)\lambda_j^{d_a}\right)$$

by (24). The first of these terms is asymptotically negligible by the same arguments as for (52), and the second by those for (38). This completes the proof.

References

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001), 'The distribution of exchange rate volatility', *Journal of the American Statistical Association* **96**, 42–55.

Christensen, B. J. and Nielsen, M. Ø. (2001), 'Semiparametric analysis of stationary fractional cointegration and the implied-realized volatility relation', *Department of Economics Working Paper 2001-04 (revised 2002), University of Aarhus* .

Christensen, B. J. and Prabhala, N. R. (1998), 'The relation between implied and realized volatility', *Journal of Financial Economics* **50**, 125–150.

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

Geweke, J. and Porter-Hudak, S. (1983), 'The estimation and application of long memory time series models', *Journal of Time Series Analysis* **4**, 221–238.

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

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

Hassler, U., Marmol, F. and Velasco, C. (2000), 'Residual log-periodogram inference for long-run relationships', *Preprint, Universidad Carlos III de Madrid* .

Lobato, I. N. (1999), 'A semiparametric two-step estimator in a multivariate long memory model', *Journal of Econometrics* **90**, 129–153.

Lobato, I. N. and Robinson, P. M. (1996), 'Averaged periodogram estimation of long memory', *Journal of Econometrics* **73**, 303–324.

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

Marinucci, D. and Robinson, P. M. (2001), 'Semiparametric fractional cointegration analysis', *Journal of Econometrics* **105**, 225–247.

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

Robinson, P. M. (1995*a*), 'Gaussian semiparametric estimation of long range dependence', *Annals of Statistics* **23**, 1630–1661.

Robinson, P. M. (1995*b*), 'Log-periodogram regression of time series with long range dependence', *Annals of Statistics* **23**, 1048–1072.

Robinson, P. M. and Marinucci, D. (1998), 'Semiparametric frequency domain analysis of fractional cointegration', *LSE Sticerd Working Paper 348* .

Robinson, P. M. and Yajima, Y. (2002), 'Determination of cointegrating rank in fractional systems', *Journal of Econometrics* **106**, 217–241.

Velasco, C. (2001), 'Gaussian semiparametric estimation of fractional cointegration', *Preprint, Universidad Carlos III de Madrid* .

Watson, M. W. (1994), Vector autoregressions and cointegration, *in* R. F. Engle and D. L. McFadden, eds, 'Handbook of Econometrics, Vol. IV', North-Holland, Amsterdam, chapter 47, pp. 2843–2915.

Table 1: Application to Implied-Realized Volatility Relation $m = 20$

Parameter	Initial	Two Step	Std. Error	$W_{d_e=0,\beta=1}$	Converged	Std. Error	$W_{d_e=0,\beta=1}$
$\hat{d}$	0.4628	0.4772	0.1023		0.4630	0.1062	
$\hat{d}_e$	0.1047	0.0655	0.1023	1.1062	< 0.0001	0.1062	0.0104
$\hat{\beta}$	0.8660	1.1173	0.1406		1.0098	0.0962	

Table 2: Application to Implied-Realized Volatility Relation $m = 27$

Parameter	Initial	Two Step	Std. Error	$W_{d_e=0,\beta=1}$	Converged	Std. Error	$W_{d_e=0,\beta=1}$
$\hat{d}$	0.4807	0.4871	0.0851		0.4866	0.0891	
$\hat{d}_e$	0.0840	0.0352	0.0851	3.7455	< 0.0001	0.0891	0.7913
$\hat{\beta}$	0.8660	1.1889	0.0999		1.0451	0.0507	

Table 3: Application to Implied-Realized Volatility Relation $m = 37$

Parameter	Initial	Two Step	Std. Error	$W_{d_e=0,\beta=1}$	Converged	Std. Error	$W_{d_e=0,\beta=1}$
$\hat{d}$	0.4527	0.4668	0.0741		0.4526	0.0767	
$\hat{d}_e$	0.0969	0.0666	0.0741	3.1547	< 0.0001	0.0767	0.1677
$\hat{\beta}$	0.8660	1.1423	0.0929		1.0267	0.0652	

Working Paper

2001-7: Jakob Roland Munch and Michael Svarer: Rent Control and Tenancy Duration.

2001-8: Morten Ø. Nielsen: Efficient Likelihood Inference in Non-stationary Univariate Models.

2001-9: Effrosyni Diamantoudi: Stable Cartels Revisited.

2001-16: Bjarne Brendstrup, Svend Hylleberg, Morten Nielsen, Lars Skipper and Lars Stentoft: Seasonality in Economic Models.

2001-17: Martin Paldam: The Economic Freedom of Asian Tigers - an essay on controversy.

2001-18: Celso Brunetti and Peter Lildholt: Range-based covariance estimation with a view to foreign exchange rates.

2002-1: Peter Jensen, Michael Rosholm and Mette Verner: A Comparison of Different Estimators for Panel Data Sample Selection Models.

2002-2: Torben M. Andersen: International Integration and the Welfare State.

2002-3: Bo Sandemann Rasmussen: Credibility, Cost of Reneging and the Choice of Fixed Exchange Rate Regime.

2002-4: Bo William Hansen and Lars Mayland Nielsen: Can Nominal Wage and Price Rigidities Be Equivalent Propagation Mechanisms? The Case of Open Economies.

2002-5: Anna Christina D'Addio and Michael Rosholm: Left-Censoring in Duration Data: Theory and Applications.

2002-6: Morten Ørregaard Nielsen: Efficient Inference in Multivariate Fractionally Integration Models.

2002-7: Morten Ørregaard Nielsen: Optimal Residual Based Tests for Fractional Cointegration and Exchange Rate Dynamics.

2002-8: Morten Ørregaard Nielsen: Local Whittle Analysis of Stationary Fractional Cointegration.