DEPARTMENT OF ECONOMICS

Working Paper

SEMIPARAMETRIC ANALYSIS OF STATIONARY
FRACTIONAL COINTEGRATION AND THE
IMPLIED-REALIZED VOLATILITY RELATION

Bent Jesper Christensen and Morten Ø. Nielsen

UNIVERSITY OF AARHUS · DENMARK

WORKING PAPER

SEMIPARAMETRIC ANALYSIS OF STATIONARY FRACTIONAL COINTEGRATION AND THE IMPLIED-REALIZED VOLATILITY RELATION

Bent Jesper Christensen and Morten Ø. Nielsen

Working Paper No. 2001-4

Semiparametric Analysis of Stationary Fractional Cointegration and the Implied-Realized Volatility Relation

Bent Jesper Christensen

School of Economics and Management

Building 322, University of Aarhus

DK-8000 Aarhus C

Denmark

E-mail: bjchristensen@econ.au.dk

Morten Ørregaard Nielsen

Department of Economics

Building 322, University of Aarhus

DK-8000 Aarhus C

Denmark

E-mail: monielsen@econ.au.dk

August 9, 2002

Abstract

We consider semiparametric frequency domain analysis of cointegration between long memory processes, i.e. fractional cointegration, allowing derivation of useful long-run relations even among stationary processes. The approach uses a degenerating part of the periodogram near the origin to form a narrow band frequency domain least squares (FDLS) estimator of the cointegrating relation, which is consistent for arbitrary short-run dynamics. Our main theoretical contribution is to derive the asymptotic distribution theory for the FDLS estimator of the cointegration vector in the stationary long memory case. The motivating example is the relation between the volatility realized in the stock market and the associated implicit volatility derived from option prices. An application to high-frequency U.S. stock index and option data is offered.

JEL classification: C14; C22; G13

Key words: Asymptotic distribution theory; Financial options; High-frequency data; Long memory; Long-run relation; Narrow band least squares

1

1 INTRODUCTION

Cointegration analysis has been one of the most active areas in the econometrics and time series literatures in the last 15 years, starting with the seminal contributions by Granger (1981) and Engle & Granger (1987). Much of this work has been prompted by applications in macroeconomics and finance. Most of the analysis has considered the $I(1) - I(0)$ type of cointegration in which linear combinations of two or more $I(1)$ variables are $I(0)$. Here, a process is labelled $I(0)$ if it is covariance stationary and has positive finite spectral density at the origin, and $I(1)$ if the once differenced series is $I(0)$. In the bivariate case, if y_t and x_t are $I(1)$ and hence in particular nonstationary (unit root) processes, but there exists a process e_t which is $I(0)$ and a fixed β such that

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

then x_t and y_t are said to be cointegrated. Thus, the nonstationary series move together in the sense that a linear combination of them is stationary and hence a common stochastic trend is shared.

The above notion of cointegration is based on the premise that the first differences of the raw series are somehow special. In particular, cointegrating relations among unit root processes (whose first differences are $I(0)$) are defined and studied. However, many economic and financial time series exhibit strong persistence without exactly possessing unit roots. The basic theory of cointegration offers no guidance as to the analysis of relations among such series. What is needed is a class of processes that is more general than $I(1)$ and still admits a criterion for linear co-movement of series. One such class is that of fractionally integrated processes. Thus, a process is $I(d)$ (fractionally integrated of order d) if its d'th difference is $I(0)$. Here, d may be any real number, i.e. $d = 0$ or $d = 1$ are special cases. For a precise statement, $x_t \in I(d)$ if

$$(1 - L)^d x_t = \varepsilon_t, \tag{2}$$

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

$$(1 - L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j - d)}{\Gamma(-d)\Gamma(j+1)} L^j, \qquad \Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt, \tag{3}$$

in the lag operator L $(Lx_t = x_{t-1})$. Following the original idea by Granger (1981), a natural generalization of the cointegration concept is to assume that the raw series are $I(d)$ and that a certain linear combination is $I(d_e)$, with d and d_e positive real numbers and $d_e < d$. Thus, the errors are of lower order of fractional integration than the levels. In this case, the series share fractionally integrated stochastic trends of different orders ($I(d)$ and $I(d_e)$), and a linear combination eliminates the largest.

Clearly, this allows the study of co-movement among persistent series much more generally than in the standard (exact) unit root based $I(1) - I(0)$ type cointegration case.

Fractional cointegration analysis has been carried out by Cheung & Lai (1993), Baillie & Bollerslev (1994), Dolado & Marmol (1996), and Dueker & Startz (1998), among others, using parametric methods. However, Robinson (1994) shows that conventional estimators, and in particular OLS, are inconsistent when the errors are fractionally integrated. He introduces a semiparametric method, which is later extended to multiple regressors by Lobato (1997) and further developed by Robinson & Marinucci (1998) (henceforth RM) and Marinucci & Robinson (2001), to cope with this issue. Their estimators are consistent for general orders of fractional integration d for the individual series and d_e for the errors in the cointegrating relation, and for arbitrary short run dynamics (which would have had to be precisely and correctly specified in the alternative parametric approaches). In addition, RM provide the limiting distribution of their estimators of the cointegrating relation for the case $d > 1/2, d_e \geq 0$.

In this paper, we study volatility series measured from both stock index levels (realized volatility) and options data (implied volatility). Financial volatility series almost universally exhibit strong serial dependence, but not of the unit root kind, see e.g. Lobato & Velasco (2000), Andersen, Bollerslev, Diebold & Ebens (2001), and Andersen, Bollerslev, Diebold & Labys (2001). Our motivating application is the study of the relation between the realized volatility computed from high-frequency index return data and the associated implied volatility backed out from the prices of options with lives (until expiration) matching the intervals for realized volatility calculation. For univariate stock market volatility series, Andersen, Bollerslev, Diebold & Ebens (2001) show that the parameter region $0 < d < 1/2$ is relevant. In the present paper, we show that indeed the region $d > 0, d_e \geq 0, d + d_e < 1/2$ is relevant for the orders of fractional integration and cointegration in the volatility application, thus extending the result of Andersen, Bollerslev, Diebold & Ebens (2001) to the multivariate case. In particular, both the volatility series (implied and realized) and the error term appear to be covariance stationary long memory processes, and the implied-realized volatility relation is one of stationary fractional cointegration. The condition $d + d_e < 1/2$ implies that the limiting distributions provided by RM do not apply. We derive the limiting distribution of the semiparametric estimator for this case. We show that the resulting distribution is normal, in contrast to RM's more complicated distributions for the different subcases in the region $d + d_e > 1/2$. This difference corresponds to what might have been conjectured from earlier results for long memory processes, e.g. Beran (1994).

Univariate realized volatility analysis has received much attention recently, particularly due to the work of Andersen, Bollerslev, Diebold & Ebens (2001) and Andersen, Bollerslev, Diebold & Labys (2001). For the multivariate case, Christensen & Prabhala (1998) studied the forecastability of future

3

realized volatility from current implied volatility. Using OLS regressions, they found evidence that implied volatility does contain significant information about future realized volatility, but that the implied forecast is downward biased. Considering the strong persistence in the volatility series, these inferences warrant further analysis. Thus, it is relevant to inquire whether implied and realized volatility move together, and in particular are cointegrated. In this case, the notion of fractional cointegration is required, since volatility although persistent does not exhibit unit root behaviour. If implied and realized volatility are indeed fractionally cointegrated, then OLS estimation of the relation between the two may be inconsistent, and the usual estimated standard errors erroneous. Hence, an actual fractional cointegration analysis is called for.

We implement the semiparametric fractional cointegration method empirically in the implied-realized volatility application, using our new limiting distribution theory for the estimator. We demonstrate that indeed implied and realized volatility are fractionally cointegrated, and that the bias in the implied volatility forecast of subsequent realized volatility is much less than suggested by the OLS analysis from the literature, and in fact negligible in some of our specifications. On the substantive side, this is evidence in favor of the forecastability of realized volatility from implied volatility. On the methodological side, the results show how careful modelling may improve inferences.

The rest of the paper is laid out as follows. In section 2 we set up the fractional cointegration model. We then show that in the stationary fractional cointegration case, i.e., the above case where the "total memory" of the model (the sum $d + d_e$ of the memory of the raw data and the cointegrating error) is less than $1/2$, the estimator of the coefficients of the cointegrating relation is asymptotically normal. Section 3 briefly reviews the relation between the volatility implied in option prices and the volatility realized in the stock market, as well as the cointegration implications. In section 4 we present our empirical results using an ultra-high frequency dataset. Section 5 concludes. The proof of the main theorem appears in the appendix.

2 STATIONARY FRACTIONAL COINTEGRATION

The stochastic process $\{x_t, t = 1, 2, ...\}$ generated by (2) has spectral density

$$f(\lambda) \sim g\lambda^{-2d} \qquad \text{as } \lambda \to 0^+, \tag{4}$$

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 die out at a hyperbolic rate, in contrast to the much faster exponential rate in the weak

4

dependence case. The most well-known model satisfying (4) is the fractional ARIMA model of Granger & Joyeux (1980) and Hosking (1981). For a modern textbook treatment of long memory processes, see e.g. Beran (1994). The parameter d determines the memory of the process. If $d > -1/2$ the process is invertible and possesses a linear (Wold) representation, and if $d < 1/2$ it is covariance stationary. If $d = 0$ the spectral density is bounded at the origin and the process has only weak dependence. Furthermore, if $d < 0$ the process is said to be anti-persistent, and has mostly negative autocorrelations, but if $d > 0$ the process has long memory. Throughout this paper, we shall be concerned with the case $0 \leq d < 1/2$, only. This interval is relevant for many applications in finance, see e.g. Lobato & Velasco (2000), Andersen, Bollerslev, Diebold & Ebens (2001), and Andersen, Bollerslev, Diebold & Labys (2001). 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. Two important semiparametric approaches have been developed, the log-periodogram estimator by Geweke & Porter-Hudak (1983) and Robinson (1995b) and the Gaussian semiparametric or local Whittle estimator by Künsch (1987) and Robinson (1995a), recently extended and refined by Lobato (1999) and Shimotsu & Phillips (2002) among others. Estimators based on fully specified parametric models are considered by Fox & Taqqu (1986) and Dahlhaus (1989) among others. The semiparametric estimators of the memory parameter assume only (4) 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. The fully parametric approaches are asymptotically efficient, using the entire sample, but are inconsistent if the parametric model is specified incorrectly, e.g., if the lag-structure of the short-term dynamics is misspecified.

A natural generalization of the standard $I(1) - I(0)$ cointegration concept is to assume that the raw data variables are $I(d)$ and that a certain linear combination is only $I(d-b)$ with $d \geq b > 0$. To be more specific, suppose the p raw data series are gathered in $z_t = (x_t', y_t)' \in I(d_1, ..., d_p)$ and that a certain linear combination of z_t is $I(d_e)$ with $d_e < \min(d_i)$. We shall then call z_t fractionally cointegrated and write $z_t \in FCI(d_1, ..., d_p; d_e)$. This would occur, for example, if $(x_t, y_t) \in I(d)$ and $e_t \in I(d_e)$ with $d > d_e \geq 0$ in the model

$$y_t = \beta' x_t + e_t. \tag{5}$$

Conceptually, an additional condition on the integration orders $d_1, ..., d_p$ is needed. In particular, it is necessary that the two largest orders of integration are equal. To extend on the previous example, suppose $(x_{1t}, x_{2t}, y_t) \in I(d_1, d_2, d_3)$ and that the variables are indexed in increasing order of magnitude of d_i. It would then clearly be necessary that $d_2 = d_3$ for x_{2t} and y_t to cointegrate. However, no

restriction is necessarily imposed on d_1. To see why, suppose $y_t - \beta_2 x_{2t} \in I(d')$, where $d' < d_2 = d_3$. That is, y_t and x_{2t} cointegrate to an $I(d')$ process. Then it would only be possible for x_{1t} to enter nontrivially in the cointegrating relation if $d' = d_1$. Otherwise, we may let $\beta_1 = 0$ and exclude x_{1t} from the equation. Thus, what is needed is a generalization of the multicointegration concept to allow the orders of integration to be different, but maintaining a structure that ensures that only variables of the same integration order cointegrate with each other.

Some authors have estimated fractionally cointegrated models. For instance, Dolado & Marmol (1996) assume an arbitrary value of d for both the raw data and the errors, and Dueker & Startz (1998) estimate a fully parametric model of the error correction type. The main drawback of fully specified parametric models is that they provide inconsistent estimates of the long-run parameters if the model is not correctly specified, e.g., if the short-run lag structure in the error correction model is incorrect. Robinson (1994) and subsequently Robinson & Marinucci (1998) and Marinucci & Robinson (2001) attempt to correct this by considering a semiparametric approach to the estimation of the cointegration equation (5). However, they only consider the asymptotic distribution theory for the case $d > 1/2$, where the raw series are nonstationary. This does not apply for the stationary long memory volatility series that we consider below. Hence, we develop the necessary theory for the case $d \in (0, 1/2)$.

2.1 The Model and Notation

In this subsection, we follow Robinson & Marinucci (1998) in setting out the model and the assumptions and present their consistency result. In the following subsection, we develop the new asymptotic distribution theory for the stationary case $d \in (0, 1/2)$. Suppose we observe the random $p-$vectors $z_t = (x_t', y_t)'$, $t = 1, \ldots, n$.

Assumption A: The vector process $\{z_t, \ t = 0, \pm 1, \ldots\}$ is covariance stationary with spectral density matrix satisfying

$$f_{zz}(\lambda) \sim \Lambda G \Lambda \quad \text{as } \lambda \to 0^+,$$

where G is a $p \times p$ real symmetric matrix whose leading $(p-1) \times (p-1)$ submatrix has full rank and

$$\Lambda = diag\left\{\lambda^{-d_1}, \ldots, \lambda^{-d_p}\right\}$$

for $d_i \in (0, 1/2)$, $i = 1, \ldots, p$. However, there exists a $(p-1)-$vector $\beta \neq 0$ and a constant $c \in (0, \infty)$ such that

$$(-\beta', 1) f_{zz}(\lambda) \begin{pmatrix} -\beta \\ 1 \end{pmatrix} = f_e(\lambda) \sim c\lambda^{-2d_e} \quad \text{as } \lambda \to 0^+,$$

with $0 \leq d_e < \min(d_i)$.

Assumption B: z_t admits the Wold representation

$$z_t = \mu_z + \sum_{j=0}^{\infty} A_j \varepsilon_{t-j}, \qquad \sum_{j=0}^{\infty} \|A_j\|^2 < \infty,$$

where the innovations ε_t satisfy

$$E\left(\varepsilon_t | \mathcal{F}_{t-1}\right) = 0, \qquad E\left(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}\right) = \Sigma, \; a.s.,$$

and $\varepsilon_t \varepsilon_t'$ are uniformly integrable. Here, $\mu_z = E(z_0)$, Σ is a constant matrix of full rank, $\mathcal{F}_t = \sigma\left(\{\varepsilon_s, s \leq t\}\right)$ is the σ-field of events generated by $\{\varepsilon_s, s \leq t\}$, and $\|\cdot\|$ is the Euclidean matrix norm.

Assumption A is a natural multivariate generalization of (4) including the multivariate fractionally integrated ARMA model as a special case, see also Robinson (1995b) and Lobato (1997). In particular, it implies that $z_{it} \in I(d_i)$ and $e_t \in I(d_e)$. Thus, Assumption A follows if $z_t \in FCI(d_1, \ldots, d_p; d_e)$. Notice that even though it is not stated as part of Assumption A, the two largest integration orders must be equal as mentioned earlier. The rank condition on G ensures no multicollinearity among the x_t variables at the origin, i.e. it restricts the x_t from cointegrating among themselves. Assumption B is a regularity condition which requires that the first and second moments of the innovations in the Wold representation of z_t are martingale difference series.

Define the Discrete Fourier Transform (DFT) of the observed vector $\{a_t, \; t = 1, \ldots, n\}$

$$w_a(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} a_t e^{it\lambda}.$$

If $\{b_t, \; t = 1, \ldots, n\}$ is another observed vector, the cross periodogram matrix between a_t and b_t is

$$I_{ab}(\lambda) = w_a(\lambda) w_b^*(\lambda) = I_{ab}^c(\lambda) + i I_{ab}^q(\lambda),$$

where the asterisk is transposed complex conjugation, and c, q indicate the co- and quadrature periodograms, respectively. We then form the discretely averaged co-periodogram

$$\hat{F}_{ab}(k, l) = \frac{2\pi}{n} \sum_{j=k}^{l} I_{ab}^c(\lambda_j), \qquad 1 \leq k \leq l \leq n-1, \tag{6}$$

for $\lambda_j = 2\pi j/n$. If $\hat{F}_{ab}$ is a vector we denote the i'th entry $\hat{F}_{ab}^{(i)}$, and if it is a matrix we denote the (i, j)'th entry $\hat{F}_{ab}^{(i,j)}$. We could also have considered a continuously averaged version of (6), but it would not enjoy the property of being invariant to mean terms.

7

With $\hat{F}$ defined as in (6) we can consider the Frequency Domain Least Squares (FDLS) estimator

$$\hat{\beta}_m = \hat{F}_{xx}^{-1}(1, m)\, \hat{F}_{xy}(1, m) \tag{7}$$

of β in the regression (5). If

$$\frac{1}{m} + \frac{m}{n} \to 0 \quad \text{as } n \to \infty, \tag{8}$$

then $\hat{\beta}_m$ is called a narrow-band FDLS estimator, since it uses only a degenerating band of frequencies around the origin. We need m to tend to infinity to gather information, but we also need to remain in a neighbourhood of zero, so m/n must tend to zero. Notice that $\hat{\beta}_{n-1}$ is the OLS estimate of β with allowance for a non-zero mean in e_t.

In this setup, RM show the following result.

Theorem 1 (Robinson and Marinucci, 1998) *Under Assumptions A, B, (8), and with $\hat{\beta}_m$ defined in (7),*

$$\hat{\beta}_{im} - \beta_i = O_p\left(\left(\frac{n}{m}\right)^{d_e - d_i}\right), \quad i = 1, \ldots, p-1, \qquad \text{as } n \to \infty.$$

Notice that under fractional cointegration $d_e < \min(d_i)$, so the estimator $\hat{\beta}_m$ is consistent for β. Furthermore, if the integration order of the raw data series is common, i.e. $d_i = d$ for all $i = 1, ..., p$, the stochastic order of the estimator varies inversely with the strength of the cointegrating relation $b = d - d_e$. In the case of different orders d_i, the estimated coefficients on the variables of highest order of integration converge fastest.

This completes the RM setup and their consistency result. As already noted, no asymptotic distribution theory has been available for the case $d_i \in (0, 1/2)$, which is relevant for our financial volatility application below. Hence, we now turn to the development of such a theory.

2.2 Limiting Distribution Theory

To derive the limiting distribution of $\hat{\beta}_m$ for $d_i \in (0, 1/2)$, we need to strengthen Assumptions A and B and introduce new assumptions similar to those in Robinson (1995b) and Lobato (1999). As RM, we find it convenient to use the notation $w_t = (x_t', e_t)'$. Henceforth, $f(\lambda) = \{f_{ij}(\lambda)\}_{i,j=1,\ldots,p}$ denotes the spectral density matrix for w_t. Note that the cross spectral density between x_{it} and e_t is $f_{ip}(\lambda)$.

Assumption A': The spectral density matrix of w_t satisfies

$$f(\lambda) \sim \Lambda G \Lambda \quad \text{as } \lambda \to 0^+.$$

In particular, there exists $\alpha \in (0, 2]$ such that

$$\left| f_{ij}(\lambda) - g_{ij}\lambda^{-d_i - d_j} \right| = O\left(\lambda^{\alpha - d_i - d_j} \right) \quad \text{as } \lambda \to 0^+, \quad i, j = 1, ..., p - 1, \tag{9}$$

where g_{ij} is the (i, j)'th element of G, which has the same properties as in Assumption A and $g_{ip} = g_{pi} = 0$, $i = 1, ..., p - 1$.

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

Assumption C: As $\lambda \to 0^+$

$$\frac{dA_i(\lambda)}{d\lambda} = O\left(\lambda^{-1} \|A_i(\lambda)\| \right)$$

for $i = 1, ..., p$, where $A_i(\lambda)$ is the $i'th$ row of $A(\lambda) = \sum_{j=0}^{\infty} A_j e^{ij\lambda}$.

Assumption D: The bandwith m satisfies

$$\frac{1}{m} + \frac{m^{1+2\alpha}}{n^{2\alpha}} \to 0 \quad \text{as } n \to \infty,$$

with α from Assumption A'.

Some comments on our assumptions are in order. Assumptions A' and C strengthen Assumption A by imposing a rate of convergence on $f(\lambda)$ and a smoothness condition similar to those used in parametric models, see e.g. Fox & Taqqu (1986) or Dahlhaus (1989). Assumption A' is satisfied with $\alpha = 2$ if, for instance, w_t is a vector fractional ARIMA process. Here, (9) with $g_{ip} = g_{pi} = 0$, $i = 1, ..., p-1$, is the central assumption of cointegration, restricting the coherencies between the cointegration residuals and other right hand side variables at the origin. The restriction is quite weak and does not require the regressors x_t and the errors e_t to be uncorrelated at any frequency, not even at the origin. Thus, our estimator allows the regressor and error terms to share the same short- and medium-term dynamics, and even allows long memory in the relation between the two. This is in stark contrast to most familiar estimation methods, such as OLS and other cointegration methods. In particular, under our assumptions, the OLS estimator of β is generally inconsistent. Note that under Assumption A' with the new notation, the condition that the two largest integration orders in z_t must be equal is automatically satisfied since $d_y = \max(d_i)$ by construction.

Assumption B' follows Robinson (1995a) and Lobato (1999) in imposing a linear structure on w_t with square summable coefficients and martingale difference innovations with finite fourth moments. It

is satisfied for instance if ε_t form an *i.i.d.* process with finite fourth moments. Under Assumption B' we can write the spectral density matrix of w_t as

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

where the asterisk is complex conjugation combined with transposition. Assumption D implies (8), so our estimator is in the narrow-band FDLS class, and if α is high the implied constraint on the bandwidth m is weak. If, e.g., w_t is a vector fractional ARIMA process such that $\alpha = 2$ we can take $m = o(n^{4/5})$.

We are now ready to derive the limiting distribution of the narrow-band FDLS estimator of the cointegrating relation in (5). The proof is in the appendix.

Theorem 2 *Under Assumptions A', B', C, D, $d_i + d_e < 1/2$, $0 \le d_e < d_i < 1/2$ for $i = 1, ..., p-1$, and with $\hat{\beta}_m$ defined in (7),*

$$\sqrt{m}\lambda_m^{d_e}\tilde{\Lambda}(\hat{\beta}_m - \beta) \xrightarrow{d} N\left(0, cH^{-1}JH^{-1}\right)$$

as $n \to \infty$, where $\tilde{\Lambda}$ is the leading $(p-1) \times (p-1)$ submatrix of Λ and H, J are defined in equations (20) and (36) in the appendix.

Note that the limiting distribution in Theorem 2 is a centered normal, unlike the non-normal distributions arising when $d_i \ge 1/2$ (Robinson & Marinucci (1998) and Marinucci & Robinson (2001)). The particular result that the asymptotic bias of $\hat{\beta}_m$ disappears is due to the key cointegration condition (9) and $g_{ip} = g_{pi} = 0$, $i = 1, ..., p-1$, in Assumption A'. This allows for straightforward inference, e.g. based on standard t-ratios (no simulation needed). Note also that in the simple two-variable case with $z_t \in I\left(d\right)$, $e_t \in I\left(d_e\right)$, and $\tilde{G} = g$, the asymptotic distribution of $\hat{\beta}_m$ in (5) is given by Theorem 2 as

$$\sqrt{m}\lambda_m^{d_e-d}\left(\hat{\beta}_m - \beta\right) \xrightarrow{d} N\left(0, \frac{c\left(1-2d\right)^2}{2g\left(1-2d-2d_e\right)}\right). \tag{11}$$

The asymptotic distribution (11) is easy to interpret and compare to well-known cases. The parameters c and g correspond to $var\left(e_t\right)$ and $var\left(x_t\right)$. Ordinary full band spectral regression yields the asymptotic variance c/g in the short memory case. In that case, Brillinger (1981, chap. 7-8) and others have also considered narrow band analysis and obtained the division by two in the asymptotic variance, as in (11). This arises since the cross-terms involving e_t vanish asymptotically in $var(\hat{F}_{xe})$ (see the proof of Theorem 2, immediately before (35)). Our Theorem 2 generalizes the analysis to allow long memory at the origin in the regressors as well as in the errors. Of course, the short memory results appear as the special case $d = d_e = 0$ of Theorem 2.

The parameters c, G, d_i, $i = 1, ..., p$, in the limiting distribution in Theorem 2 can be replaced by consistent estimates. Many such estimates are available, but the Gaussian semiparametric estimator of Robinson (1995a) and Lobato (1999) has particularly simple asymptotic properties, and is also employed by Marinucci & Robinson (2001) to estimate the fractional integration orders in their analysis.

To assess the strength of the cointegrating relation, one could calculate the observed error process $\hat{e}_t = y_t - \hat{\beta}'_m x_t$, $t = 1, ..., n$, and estimate $\hat{b} = \hat{d} - \hat{d}_{\hat{e}}$ (assuming a common integration order $d_i = d$ for the raw data series) using e.g. the methods mentioned above. Formally, Velasco (2001) shows that Robinson's (1995a) Gaussian semiparametric estimator applies for residuals under additional regularity and bandwidth conditions. In any case, this provides an informal diagnostic.

The condition that $d_i + d_e < 1/2$ for $i = 1, ..., p$ in Theorem 2 includes many relevant cases, both empirically and from theoretical models, e.g. the case when the raw data series are stationary with long memory and the error process - the deviations from equilibrium - has only short memory (or is actually white noise) due to rational expectations.

In the case where $d_i < 1/2 \leq d_i + d_e < 1$ for some i, it is conjectured that $\hat{\beta}_m$ converges to a function of the Rosenblatt process like the scalar averaged periodogram estimator in Lobato & Robinson (1996). It is also conjectured that even though $\hat{\beta}_m$ is asymptotically non-normal in this case, it belongs to the Local Asymptotically Mixed Normal (LAMN) family which has several important implications, one of them being that Wald, LM, and LR tests on $\hat{\beta}_m$ are asymptotically χ^2 distributed.

To sum up, Theorem 2 provides the necessary asymptotic distribution theory for the stationary fractional cointegration model, thus complementing the consistency result of Robinson (1994) and Robinson & Marinucci (1998). We use this in the empirical application below.

3   THE IMPLIED-REALIZED VOLATILITY RELATION

Our motivating application of stationary fractional cointegration analysis is the relation between the volatility implied in option prices and the subsequent realized return volatility of the underlying asset.

If option market participants are rational and markets are efficient, then the price of a financial option should reflect all publicly available information about future return volatility of the underlying asset over the life of the option. Empirical analysis of this hypothesis has typically employed the option pricing formula of Black & Scholes (1973) and Merton (1973) - henceforth the BSM formula. According to this, the fair (arbitrage free) price of a European call option with τ periods to expiration and strike

price k is given by

$$c\left(s, k, \tau, r, \sigma\right) = s\Phi\left(d\right) - e^{-r\tau}k\Phi\left(d - \sigma\sqrt{\tau}\right), \tag{12}$$

$$d = \frac{\ln\left(s/k\right) + \left(r + \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}},$$

where s is the price of the underlying asset, r is the riskless interest rate, Φ is the standard normal c.d.f., and σ is the return volatility of the underlying asset through expiration of the option (τ periods hence).

Given an observation c of the price of the option, the implied volatility σ_{IV} may be determined by inverting (12), i.e. solving the nonlinear equation

$$c = c\left(s, k, \tau, r, \sigma_{IV}\right) \tag{13}$$

numerically with respect to σ_{IV}, for given data on s, k, τ, and r. If this is done every period t, a time series $\sigma_{IV,t}$ results. Each implied volatility $\sigma_{IV,t}$ may now be compared to the actually realized return volatility of the underlying asset, from the time t where $\sigma_{IV,t}$ is calculated and until expiration of the option at $t + \tau$. Here, realized volatility is simply the sample standard deviation $\sigma_{RV,t}$ of the realized return from t to $t + \tau$.

Christensen & Prabhala (1998) (henceforth CP) considered the regression specification

$$y_t = \alpha + \beta x_t + e_t, \tag{14}$$

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. 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. The options were at-the-money ($k_t = s_t$) one-month ($\tau = 1/12$) OEX contracts. CP also presented results without the log transform, and the difference was negligible. Basic OLS regression in (14) produced a β-estimate that was significantly greater than zero and less than unity.

If volatility is persistent and, indeed, fractionally integrated, as empirical litterature suggests (Andersen, Bollerslev, Diebold & Ebens (2001) find fractional integration with d at about $0.35 - 0.4$), whereas the forecasting error e_t in (14) is serially uncorrelated or possesses only short memory, then x_t and y_t are fractionally cointegrated. This is in fact what our empirical results below indicate. Thus, all inferences from OLS must be called into question. In particular, even though OLS is super-consistent for β in the standard $I\left(1\right) - I\left(0\right)$ cointegration case, it is generally inconsistent in the case of stationary fractional cointegration (Robinson (1994), Robinson & Marinucci (1998)). The narrow band analysis is consistent, and has also been applied recently by Bandi & Perron (2002) to the implied-realized volatility

relation. However, due to the lack of available asymptotic distribution theory, they rely on subsampling for inference.

In the present paper, we reexamine the volatility forecasting hypothesis within the relevant stationary fractional cointegration framework, invoking our new distribution theory for inference. Long memory warrants an examination of the relation between implied and realized volatility at the long-run frequency. The condition $\beta = 1$ is then a long-run unbiasedness hypothesis. This is a reinterpretation relative to CP, who considered the stronger conditions of unbiasedness at all frequencies and market efficiency. They ascribed findings of β less than unity to correlation between regressors and errors, induced e.g. by measurement errors or possible misspecification of the BSM model. By focusing on a diminishing band of low frequencies, the role of such correlation problems vanish asymptotically, provided the spectral density of the measurement errors is dominated by that of the observed variables, i.e. is of lower order of fractional integration.

4 DATA AND EMPIRICAL RESULTS

4.1 Data

We sample options data from the Berkeley Options Data Base (BODB) (see the BODB User's Guide or Rubinstein & Vijh (1987) for a description). We use all quotes from January 1, 1988, to December 31, 1995, for options on the S&P500 index (SPX options). The SPX options are traded frequently and heavily. When using the BSM formula (12), the SPX options have the advantage over the OEX contracts considered by CP that they are European style, as assumed in the formula, i.e. there is no early exercise. Quotes are revised frequently due to the heavy trading and are used here since quotes data are expected to be more reliable than trading prices. Quotes data are recorded automatically and instantaneously when quotes are revised, whereas trading prices are recorded manually with a time lag. This could bias results for high-frequency data. Finally, our sampling period starts after the October 1987 stock market crash, since CP documented a regime shift in the implied-realized volatility relation around the time of the crash. Volatility was elevated for a period following the crash, and hence we exclude the remainder of the year 1987.

From the high-frequency options data, a 5-minute return series for the underlying (the S&P500 index) is constructed. Thus, there is a total of 8,189,617 quotes between January 1, 1988 and December 31, 1995, and we select the quotes closest to the 5-minute points during each trading day. These quotes are associated with time stamps 9:00 AM, 9:05 AM, ..., 2:55 PM, 3:00 PM. This results in a series of 147,022 observations. The average absolute error in matching the desired time stamps is 25.18 seconds.

The raw mean of the timing errors is 7.87 seconds, with a standard error of the mean 0.6532 seconds. In the BODB, a simultaneous price of the underlying is recorded for each quoted option price. We use these underlying index levels to calculate our 5-minute return series.

The high-frequency (5-minute) return series forms the basis for realized volatility calculations. In particular, we choose a one-week interval and calculate the realized variance

$$\sigma^2_{RV,t} = \frac{1}{K-1} \sum_{k=1}^{K} (r_{t,k} - \bar{r}_t)^2, \tag{15}$$

where $r_{t,k}$ are the 5-minute annualized returns in week t, and $\bar{r}_t$ is the weekly average. The realized volatility is the square-root of $\sigma^2_{RV,t}$. In practice, we start Monday morning at 10:00 AM and end Friday at 3:00 PM to avoid effects of irregular behavior from the Friday close to the Monday opening. Thus, each estimate is based on up to $K = 348$ high-frequency return observations. For comparison, Andersen, Bollerslev, Diebold & Ebens (2001) use 5-minute returns to form daily realized volatilities and have 79 returns in each.

In addition, we need implied volatilities. Here, we use the Monday 10:00 AM quote (according to the above definition) for the call of shortest maturity and closest to the money. Since options expire at 3:00 PM on the Friday immediately preceding the third Saturday of each month, the sampled options will have $4\frac{5}{6}$, $9\frac{5}{6}$, $14\frac{5}{6}$, $19\frac{5}{6}$, or $24\frac{5}{6}$ trading days to expiration, except that business holidays may reduce each of these figures slightly. From each sampled quote, an implied volatility is backed out using the BSM formula (12), as described in the previous section. We correct for dividends on the S&P500 index as described in Hull (1997, p. 263). Dividend yield data are obtained from Datastream. We use two different measures of time to expiration to reflect that calendar days are relevant for interest and dividends, and trading days for volatilities, following Hull (1997, p. 248). This results in a weekly implied volatility series $\tilde{\sigma}_{IV,t}$. However, if the implied volatility $\tilde{\sigma}_{IV,t}$ in week t is from an option with one week ($d_0 = 4\frac{5}{6}$ trading days) to expiration then $\tilde{\sigma}_{IV,t-i}$ corresponds to $i+1$ weeks ($d_i = 5i + 4\frac{5}{6}$ days) to expiration, $i = 0, 1, 2, 3$, and in some cases also for $i = 4$. In the other cases $\tilde{\sigma}_{IV,t-4}$ is again from a one week option, and which case applies depends on when the third Saturday in the relevant month occurs. 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 (length d_0) intervals by the formula

$$\sigma^2_{IV,t-i} = \frac{1}{d_i - d_{i-1}} \left(d_i \cdot \tilde{\sigma}^2_{IV,t-i} - d_{i-1} \cdot \tilde{\sigma}^2_{IV,t-i+1} \right), \tag{16}$$

starting with $\sigma_{IV,t} = \tilde{\sigma}_{IV,t}$ for t corresponding to a one week option and applying (16) for $i = 1, 2, 3$ and if applicable also $i = 4$. Here, (16) is an identity for the associated realized volatilities covering interval

lengths d_0, d_i, and d_{i-1}. This identity becomes an approximation for implied (as opposed to realized) volatilities. In practice, if the RHS of (16) turns out negative, we let $\sigma_{IV,t-i} = \sigma_{IV,t-i-1}$ for that week. This occurs in 30 of the 417 weeks in our sample, or about once every 14 weeks. Thus, due to the approximations involved, our results below on the forecasting performance of weekly implied volatility are conservative estimates based on the sample at hand. As noted earlier, the effect of measurement errors vanishes asymptotically.

4.2 Empirical Results

Summary statistics for the two volatility series constructed in Section 4.1 appear in Table 1. Each of the series is of length 417 weeks, covering the interval January 1988 through December 1995. From Table 1, implied volatility has a higher mean than realized volatility. In the case of American style options, such a difference would be at least partly attributed to the early exercise premium embedded in implied volatility. In particular, the BSM formula (12) does not correct for the possibility of early exercise. However, we have deliberately chosen the S&P500 (SPX) options because they are of European style, so there is no early exercise premium issue. Hence, the difference in mean volatility in our case must be attributed to excess hedging costs associated with replicating the options from the underlying (assumed zero in (12)).

Table 1 about here

The next line in Table 1 shows that implied volatility also has higher variance than realized volatility. This is inconsistent with the notion that implied volatility is the market's rational forecast of subsequent index volatility, i.e., a conditional expectation (see Section 3). In effect, we have only a noisy estimate of true implied volatility, due to the approximation in (16), among others, and as already noted, the assessment of the implied-realized volatility relation is conservative in any given sample, even if it is precise in large samples.

The third line in Table 1 shows that there is very little skewness in realized and implied volatility. Realized volatility is slightly positively skewed, and implied volatility is slightly negatively skewed, but the magnitudes are negligible. The next line in Table 1 shows that realized volatility is not very leptokurtic, consistent with the findings in Andersen, Bollerslev, Diebold & Ebens (2001). However, the table also shows that implied volatility in fact is somewhat leptokurtic. This difference is an interesting addition to the picture of the properties of volatility which has not been noted in the literature. Thus,

our results so far complement those of Andersen, Bollerslev, Diebold & Ebens (2001) who find that log volatilities are indeed Gaussian.

Figures 1-2 about here

The sample autocorrelation functions for the realized and implied volatility series are exhibited in Figures 1 and 2. For both series, the decay is very slow, with significant autocorrelations even beyond the 50th lag. Even when autocorrelations drop below the upper 95% confidence limit shown in the figures, they all remain positive. This is consistent with the hypothesis of long memory in volatility and with the results of Andersen, Bollerslev, Diebold & Ebens (2001).

Next, in Table 2 we turn to the analysis of the memory properties for each individual volatility series. First, the impression of highly significant autocorrelation functions for both realized and implied volatility is confirmed by the Box-Pierce statistics with 4 or 24 lags (roughly one and six months) in the first two lines of the table. The second part of the table shows Gaussian semiparametric (henceforth GSP, see Robinson (1995a)) estimates of the fractional integration parameter d for each series, and for different choices of the bandwidth parameter m.

Figures 3-4 about here

Figures 3 and 4 show the log periodograms for both series plotted against the log Fourier frequencies and both exhibit negative linear trends at the long frequencies (i.e. the periodograms have peaks at the long frequencies), thus reinforcing the impression of long memory series. The chosen bandwidths are $n^{0.5} = 20$, $n^{0.6} = 37$, and $n^{0.7} = 68$, corresponding roughly to the log frequencies -1 to 0, up to which the negative linear trend in Figures 3 and 4 is most pronounced. Returning to Table 2, all the estimates are below $1/2$, but significantly greater than 0 (asymptotic standard errors in parentheses). This suggests that both realized and implied volatility are covariance-stationary long memory series.

Table 2 about here

To be sure that the long memory evident in the volatility series does not in fact reflect a unit root, we also in the last part of Table 2 exhibit standard augmented Dickey-Fuller unit root tests with 0 and 4 lags. These tests complement the GSP estimates and show clearly that the volatility series are not unit root processes.

The results so far are consistent with the notion that realized and implied volatility both are driven by stationary but fractionally integrated series. The interesting question is how closely they move together.

Table 3 about here

16

Under the long-run unbiasedness hypothesis, we would expect the series to follow each other closely. In Table 3, we turn to the analysis of the potential stationary fractional cointegration relation between the two series. The first line of the table shows the results of the OLS regression

$$y_t = \alpha + \beta x_t + e_t \tag{17}$$

of realized on implied volatility. The parameter of interest, β, is estimated to be 0.38, but viewing OLS as a special case of FDLS, we have the bandwidth parameter $m = n - 1$ ($n = 417$ is the number of time series observations), and this is too much for calculating standard errors under the maintained hypothesis (based on Table 2) of long memory in the individual volatility series. The point estimate, however, is similar to those in the literature, and suggests that implied volatility is a downward biased forecast of realized volatility. Of course, the properties of OLS are called into question when the possibility of fractional cointegration is recognized. Note the high estimated order of integration of the OLS residuals (0.31 to 0.39, depending on bandwidth), i.e. the residuals possess long memory.

Turning to lower bandwidth parameters m for the estimation of the parameter of interest, β, we find much larger point estimates. Following Robinson & Marinucci (1998) (see also Marinucci & Robinson (2001)), we actually place considerable weight on the estimates resulting from low values of m, such as $m = 3$ or $m = 6$, and only consider estimation with up to $m = 15$ (RM consider $m = 3, 4, 6$ for their sample sizes of $n = 116$ and $n = 138$). For example, for $m = 15$ we find $\hat{\beta}$ in excess of 0.85, i.e. more than twice the value from OLS, and the estimate is significantly greater than zero. Here, the asymptotic standard errors of $\hat{\beta}_m$ are calculated using the new asymptotic distribution result in Theorem 2, with the results from Table 2 (we have chosen bandwidth 68, since the results are quite close and this choice yields results closest to those in the literature) for d. For $m = 15$ or less, $\hat{\beta}_m$ is insignificantly less than unity, whether using $m_e = 20$, 37, or 68 for the bandwidth parameter to estimate d_e (this enters the formula from Theorem 2 for the standard error of $\hat{\beta}_m$).

From the residual analysis, we cannot reject the null hypothesis 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, $d_e < d$ and $d_e + d < 1/2$. In fact, our results are consistent with the even stronger relation that $d_e = 0$. Thus, d_e is insignificant in Table 3 for $m = 15$ and lower, whether using bandwidth $m_e = 20$, 37 or 68 for estimating d_e.

Figures 5-6 about here

Here we have chosen the same bandwidths as in the estimation of the memory parameter of the raw data. The log periodogram vs. log Fourier frequency plots do not give any clear guidance to the

17

choice of bandwidth for the residuals. For example, the plots using $m = 3$ and $m = 15$ (to estimate β and generate the residuals) are given in Figures 5 and 6. The results indicate that the cointegration error process exhibits only short memory, so that all long memory properties in volatility are common features for implied and realized volatility.

Considering different bandwidths m for the estimation of β, we note a monotonicity, with a tendency towards higher $\hat{\beta}_m$ for lower m. Thus, the more we focus the analysis on the long frequencies, the less the data that we use to extract information about the long term relation between the series is contaminated by short term noise, which could include measurement errors. For $m = 3$, we find $\hat{\beta} = 0.89$ and with a rather narrow confidence band which contains unity. Clearly, these results leave the possibility of $\beta = 1$ very likely, consistent with the notion that implied volatility from option prices is an unbiased forecast of the subsequently realized index volatility in the long run. This is obviously in stark contrast to the raw OLS results in the first line of the table, with $\hat{\beta} = 0.38$. It should be noted in this connection that the usual (inconsistent) OLS standard error for $\hat{\beta}$ (not reported in the table) is 0.033. Clearly, inferences on volatility relations may be heavily misleading if the possibility of stationary fractional cointegration is ignored.

5 CONCLUSION

Option market efficiency implies a coefficient of unity in the implied-realized volatility relation. The ordinary regression estimate is less than one-half. We show that the implied and realized volatility series are stationary fractionally cointegrated. When accounting for this, the revised estimate is more than twice as large as before, and insignificantly less than unity. This shows the need to account for long-run relations even among stationary series. In effect, the evidence is in favor of a long-run unbiasedness hypothesis, which is a weaker condition than market efficiency.

We employ the semiparametric frequency domain least squares estimator by Robinson (1994). This estimator has been used by Robinson & Marinucci (1998) and Marinucci & Robinson (2001) in the case of nonstationary series. 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, an asymptotic distribution theory for the estimator of the cointegrating relation has been missing. We have provided and proved the necessary new theory in this paper. It is noteworthy and useful that our asymptotic distribution is normal, in contrast to the more complicated distributions arising in the nonstationary case.

ACKNOWLEDGEMENTS

APPENDIX: PROOF OF THEOREM 2

We now turn to the proof of Theorem 2. Following Robinson (1995a) and Lobato (1999), the proof of asymptotic normality of $\hat{F}_{xe} = \hat{F}_{xe}(1, m)$ first approximates the cross periodogram between x_{it} and e_t by $A_i(\lambda) J(\lambda) A_p^*(\lambda)$, where $J(\lambda)$ is the periodogram of ε_t by Assumption B', and then invokes a central limit theorem for martingale difference arrays.

Proof of Theorem 2. We want to examine each of the terms in

$$
\begin{aligned}
\sqrt{m}\lambda_m^{d_e}\tilde{\Lambda}(\hat{\beta}_m - \beta) &= \sqrt{m}\lambda_m^{d_e}\tilde{\Lambda}\hat{F}_{xx}^{-1}(1, m)\hat{F}_{xe}(1, m) \\
&= \left\{\lambda_m\tilde{\Lambda}\hat{F}_{xx}^{-1}(1, m)\tilde{\Lambda}\right\}\left\{\sqrt{m}\lambda_m^{d_e-1}\tilde{\Lambda}^{-1}\hat{F}_{xe}(1, m)\right\}.
\end{aligned}
\tag{18}
$$

Using linearity of $\operatorname{Re}(\cdot)$ and Theorem 1 of Lobato (1997), which is implied by our assumptions, we conclude that

$$
\hat{F}_{ik} - F_{ik} = o_p\left(\sqrt{F_{ii}}\sqrt{F_{kk}}\right), \quad 1 \le i, k \le p - 1.
$$

Thus, the first term in (18) is

$$
\lambda_m\tilde{\Lambda}\hat{F}_{xx}^{-1}(1, m)\tilde{\Lambda} \xrightarrow{p} \lambda_m\tilde{\Lambda}\tilde{\Lambda}^{-1}\lambda_m^{-1}H^{-1}\tilde{\Lambda}^{-1}\tilde{\Lambda} = H^{-1}
\tag{19}
$$

by the Continuous Mapping Theorem, where

$$
H_{ik} = \frac{g_{ik}}{1 - d_i - d_k}, \quad 1 \le i, k \le p - 1.
\tag{20}
$$

Note that the leading $(p-1) \times (p-1)$ submatrix of G (and thus H) is invertible by Assumption A.

For the remaining part of the proof we look at the convergence in distribution of

$$
\sqrt{m}\lambda_m^{d_e-1}\tilde{\Lambda}^{-1}\hat{F}_{xe}(1, m).
$$

Applying the Cramèr-Wold device, we need to examine (η is an arbitrary $(p-1) \times 1$ vector)

$$\eta' \sqrt{m} \lambda_m^{d_e-1} \tilde{\Lambda}^{-1} \hat{F}_{xe}(1,m)$$

$$= \sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_i+d_e-1} \hat{F}_{ip}(\lambda_m)$$

$$= \sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_i+d_e-1} \frac{2\pi}{n} \sum_{j=1}^{m} \text{Re}(I_{ip}(\lambda_j))$$

$$= \sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_i+d_e-1} \frac{2\pi}{n} \sum_{j=1}^{m} \text{Re}(I_{ip}(\lambda_j) - A_i(\lambda_j) J(\lambda_j) A_p^*(\lambda_j)) \tag{21}$$

$$+ \sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_i+d_e-1} \frac{2\pi}{n} \sum_{j=1}^{m} \text{Re}(A_i(\lambda_j) J(\lambda_j) A_p^*(\lambda_j)). \tag{22}$$

By $(C.2)$ of Lobato (1999), which is implied by our assumptions, the first term, i.e. expression (21), is

$$(21) = O_p\left(\sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{-1} \frac{1}{n}\left[m^{1/3}(\log m)^{2/3} + \log m + \frac{\sqrt{m}}{n^{1/4}}\right]\right)$$

$$= O_p\left(\sum_{i=1}^{p-1} \eta_i \frac{1}{\sqrt{m}}\left[m^{1/3}(\log m)^{2/3} + \log m + \frac{\sqrt{m}}{n^{1/4}}\right]\right)$$

$$= O_p\left(\frac{(\log m)^{2/3}}{m^{1/6}} + \frac{\log m}{\sqrt{m}} + \frac{1}{n^{1/4}}\right) \to_p 0.$$

The second term above, expression (22), can be written

$$(22) = \sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_i+d_e-1} \frac{2\pi}{n} \sum_{j=1}^{m} \text{Re}\left(A_i(\lambda_j) \frac{1}{2\pi n}\left|\sum_{t=1}^{n} \varepsilon_t e^{it\lambda_j}\right|^2 A_p^*(\lambda_j)\right)$$

$$= \sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_i+d_e-1} \frac{2\pi}{n} \sum_{j=1}^{m} \text{Re}\left(A_i(\lambda_j) \frac{1}{2\pi n} \sum_{t=1}^{n} \varepsilon_t \varepsilon_t' A_p^*(\lambda_j)\right) \tag{23}$$

$$+ \sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_i+d_e-1} \frac{2\pi}{n} \sum_{j=1}^{m} \text{Re}\left(A_i(\lambda_j) \frac{1}{2\pi n} \sum_{t=1}^{n} \sum_{s \neq t} \varepsilon_t \varepsilon_s' e^{i(t-s)\lambda_j} A_p^*(\lambda_j)\right). \tag{24}$$

Now write (23) as

$$(23) = \sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_i+d_e-1} \frac{1}{n} \sum_{j=1}^{m} \text{Re}(A_i(\lambda_j) A_p^*(\lambda_j)) \tag{25}$$

$$+ \sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_i+d_e-1} \frac{1}{n} \sum_{j=1}^{m} \text{Re}\left(A_i(\lambda_j)\left(\frac{1}{n} \sum_{t=1}^{n} \varepsilon_t \varepsilon_t' - I_p\right) A_p^*(\lambda_j)\right), \tag{26}$$

where (25) is bounded by

$$\sup_i O\left(\sqrt{m}\lambda_m^{d_i+d_e-1}\frac{1}{n}\sum_{j=1}^m |f_{ip}(\lambda_j)|\right) = \sup_i O\left(\sqrt{m}\lambda_m^{d_i+d_e-1}\frac{m}{n}\lambda_m^{\alpha-d_i-d_e}\right)$$

$$= O\left(\frac{m^{1+2\alpha}}{n^{2\alpha}}\right) \to 0$$

by Assumptions A' and D. For (26) we use the fact that $\varepsilon_t\varepsilon_t'-I_p$ is a martingale difference sequence with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{Z}}$, $\mathcal{F}_t = \sigma(\{\varepsilon_s, s\le t\})$, such that in particular $n^{-1}\sum_{t=1}^n \varepsilon_t\varepsilon_t'-I_p = o_p(1)$. Then (26) is bounded by

$$\sup_i o_p\left(\sqrt{m}\lambda_m^{d_i+d_e-1}\frac{1}{n}\sum_{j=1}^m |f_{ip}(\lambda_j)|\right) = o_p\left(\frac{m^{1+2\alpha}}{n^{2\alpha}}\right) \to_p 0$$

by Assumptions A' and D.

We return to (24),

$$\sum_{i=1}^{p-1}\eta_i\sqrt{m}\lambda_m^{d_i+d_e-1}\frac{1}{n^2}\sum_{j=1}^m \text{Re}\left(A_i(\lambda_j)\sum_{t=1}^n\sum_{s\neq t}\varepsilon_t\varepsilon_s' e^{i(t-s)\lambda_j}A_p^*(\lambda_j)\right)$$

$$= \sum_{t=1}^n\varepsilon_t'\sum_{s\neq t}\sum_{i=1}^{p-1}\eta_i\frac{\sqrt{m}}{n^2}\lambda_m^{d_i+d_e-1}\sum_{j=1}^m \text{Re}\left(A_i'(\lambda_j)e^{i(t-s)\lambda_j}\bar{A}_p(\lambda_j)\right)\varepsilon_s$$

$$= \sum_{t=1}^n\varepsilon_t'\sum_{s=1}^{t-1}c_{t-s,n}\varepsilon_s,$$

defining

$$c_{tn} = \frac{1}{2\pi n\sqrt{m}}\sum_{j=1}^m \theta_j\cos(t\lambda_j),$$

$$\theta_j = \sum_{i=1}^{p-1}\eta_i\lambda_m^{d_i+d_e}\text{Re}\left(A_i'(\lambda_j)\bar{A}_p(\lambda_j) + A_p'(\lambda_j)\bar{A}_i(\lambda_j)\right).$$

So, $z_{tn} = \varepsilon_t'\sum_{s=1}^{t-1}c_{t-s,n}\varepsilon_s$ is a martingale difference array with respect to $(\mathcal{F}_t)_{t\in\mathbb{Z}}$, and we can apply the CLT if

$$\sum_{t=1}^n E\left(z_{tn}^2\big|\mathcal{F}_{t-1}\right) - \sum_{i=1}^{p-1}\sum_{k=1}^{p-1}\eta_i\eta_k cJ_{ik} \to_p 0, \tag{27}$$

$$\sum_{t=1}^n E\left(z_{tn}^2 1\left(|z_{tn}| > \delta\right)\right) \to 0,\ \delta > 0, \tag{28}$$

see Hall & Heyde (1980, chap. 3.2). A sufficient condition for (28) is

$$\sum_{t=1}^n E\left(z_{tn}^4\right) \to 0. \tag{29}$$

21

First, we show (27),

$$\sum_{t=1}^{n} E\left(z_{tn}^2 \big| \mathcal{F}_{t-1}\right) = \sum_{t=1}^{n} E\left(\sum_{s=1}^{t-1}\sum_{r=1}^{t-1} \varepsilon_s' c_{t-s,n}' \varepsilon_t \varepsilon_t' c_{t-r,n} \varepsilon_r \bigg| \mathcal{F}_{t-1}\right)$$

$$= \sum_{t=1}^{n}\sum_{s=1}^{t-1} \varepsilon_s' c_{t-s,n}' c_{t-s,n} \varepsilon_s \tag{30}$$

$$+ \sum_{t=1}^{n}\sum_{s=1}^{t-1}\sum_{r\neq s} \varepsilon_s' c_{t-s,n}' c_{t-r,n} \varepsilon_r. \tag{31}$$

By $(D.10)$ and $(D.11)$ of Lobato (1999), (31) has mean zero and variance

$$O\left(n\left(\sum_{s=1}^{n}\|c_{sn}\|^2\right)^2 + \sum_{t=3}^{n}\sum_{u=2}^{t-1}\left(\sum_{s=1}^{u-1}\|c_{u-s,n}\|^2 \sum_{s=1}^{u-1}\|c_{t-s,n}\|^2\right)\right). \tag{32}$$

Since $\|\theta_j\| = O(1)$ by construction,

$$\|c_{sn}\| = O\left(\frac{1}{n\sqrt{m}}\sum_{j=1}^{m}\|\theta_j\|\right) = O\left(\frac{\sqrt{m}}{n}\right),$$

and since $\sum_{j=1}^{k}|\cos(s\lambda_j)| = O(n/s)$ another bound is

$$\|c_{sn}\| = O\left(\frac{1}{n\sqrt{m}}\sum_{j=1}^{m}\|\theta_j\||\cos(s\lambda_j)|\right) = O\left(\frac{1}{s\sqrt{m}}\right).$$

The latter bound for $\|c_{sn}\|$ is better when $s > n/m$. Thus, we find that

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

$$= O\left(\frac{n}{m}\left(\frac{\sqrt{m}}{n}\right)^2 + \frac{1}{m}\sum_{s=[n/m]+1}^{n}s^{-2}\right)$$

$$= O\left(n^{-1}\right),$$

implying that the first term of (32) is $O\left(n^{-1}\right)$. The second term of (32) is bounded by

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

see Robinson (1995a, pp. 1646-1647) or Lobato (1999, p. 151). The summand in the last sum is $O\left(n^{-2}sm + (sm)^{-1}\right)$. Choosing the first bound when $s \leq \left[n/m^{2/3}\right]$, the last sum is

$$O\left(\sum_{s=1}^{[n/m^{2/3}]}\frac{sm}{n^2} + \sum_{s=[n/m^{2/3}]+1}^{[n/2]}\frac{1}{sm}\right) = O\left(\frac{1}{m^{1/3}}\right)$$

22

and $(32) = O\left(n^{-1} + m^{-1/3}\right)$.

We still need to show that the mean of (30) is asymptotically equivalent to $\sum_{i=1}^{p-1} \sum_{k=1}^{p-1} \eta_i \eta_k c J_{ik}$. Thus,

$$
\begin{aligned}
E(30) &= \sum_{t=1}^{n} \sum_{s=1}^{t-1} E\left(\mathrm{tr}\left(c'_{t-s,n} c_{t-s,n} \varepsilon_s \varepsilon'_s\right)\right) \\
&= \sum_{t=1}^{n} \sum_{s=1}^{t-1} \mathrm{tr}\left(c'_{t-s,n} c_{t-s,n}\right) \\
&= \sum_{t=1}^{n} \sum_{s=1}^{t-1} \sum_{j=1}^{m} \sum_{k=1}^{m} \frac{1}{4\pi^2 n^2 m} \mathrm{tr}\left(\theta'_j \theta_k\right) \cos\left((t-s)\lambda_j\right) \cos\left((t-s)\lambda_k\right) \\
&= \sum_{t=1}^{n} \sum_{s=1}^{t-1} \sum_{j=1}^{m} \frac{1}{4\pi^2 n^2 m} \mathrm{tr}\left(\theta'_j \theta_j\right) \cos^2\left((t-s)\lambda_j\right) \qquad (33) \\
&\quad + \sum_{t=1}^{n} \sum_{s=1}^{t-1} \sum_{j=1}^{m} \sum_{k \neq j}^{m} \frac{1}{4\pi^2 n^2 m} \mathrm{tr}\left(\theta'_j \theta_k\right) \cos\left((t-s)\lambda_j\right) \cos\left((t-s)\lambda_k\right). \qquad (34)
\end{aligned}
$$

Notice that, since $\|\theta_j\| = O(1)$,

$$
(34) = O\left(\sum_{t=1}^{n} \sum_{s=1}^{t-1} \sum_{j=1}^{m} \sum_{k \neq j}^{m} \frac{1}{n^2 m} \cos\left((t-s)\lambda_j\right) \cos\left((t-s)\lambda_k\right) \right)
$$

and, using that $\sum_{t=1}^{n} \sum_{s=1}^{t-1} \cos\left((t-s)\lambda_j\right) \cos\left((t-s)\lambda_k\right) = -n/2$, we bound (34) by $O\left(\sum_{j=1}^{m} \sum_{k \neq j}^{m} \left(n^2 m\right)^{-1} n\right) = O(m/n)$. Now,

$$
\begin{aligned}
\mathrm{tr}\left(\theta'_j \theta_j\right) &= \mathrm{tr}\left(\sum_{i=1}^{p-1} \sum_{k=1}^{p-1} \eta_i \eta_k \lambda_m^{d_i + d_k + 2d_e} \mathrm{Re}\left(A_p^*(\lambda_j) A_i(\lambda_j) + A_i^*(\lambda_j) A_p(\lambda_j)\right) \mathrm{Re}\left(A_k'(\lambda_j) \bar{A}_p(\lambda_j) + A_p'(\lambda_j) \bar{A}_k(\lambda_j)\right) \right) \\
&= \sum_{i=1}^{p-1} \sum_{k=1}^{p-1} \eta_i \eta_k \lambda_m^{d_i + d_k + 2d_e} 4\pi^2 \left(f_{pp}(\lambda_j) f_{ik}(\lambda_j) + f_{ip}(\lambda_j) f_{kp}(\lambda_j) + f_{pk}(\lambda_j) f_{pi}(\lambda_j) + f_{pp}(\lambda_j) f_{ki}(\lambda_j)\right)
\end{aligned}
$$

by definition of $f(\lambda)$. The second and third terms are of smaller order than the first and fourth terms by Assumption A', so (33) reduces to

$$
\sum_{t=1}^{n} \sum_{s=1}^{t-1} \sum_{j=1}^{m} \frac{1}{n^2 m} \sum_{i=1}^{p-1} \sum_{k=1}^{p-1} \eta_i \eta_k \lambda_m^{d_i + d_k + 2d_e} \left(f_{pp}(\lambda_j) f_{ik}(\lambda_j) + f_{pp}(\lambda_j) f_{ki}(\lambda_j)\right) \cos^2\left((t-s)\lambda_j\right). \qquad (35)
$$

Notice that

$$
\begin{aligned}
\frac{2\pi}{n} \sum_{j=1}^{m} \left(f_{pp}(\lambda_j) f_{ik}(\lambda_j) + f_{pp}(\lambda_j) f_{ki}(\lambda_j)\right) &\sim \int_{0}^{\lambda_m} \left(f_{pp}(\lambda) f_{ik}(\lambda) + f_{pp}(\lambda) f_{ki}(\lambda)\right) d\lambda \\
&\sim \int_{0}^{\lambda_m} \left(g_{ik} + g_{ki}\right) c \lambda^{-d_i - d_k - 2d_e} d\lambda \\
&= 2 g_{ik} c \frac{\lambda_m^{1 - d_i - d_k - 2d_e}}{1 - d_i - d_k - 2d_e},
\end{aligned}
$$

23

so that we can rewrite (35) as

$$(35) \quad \sim \quad \sum_{i=1}^{p-1}\sum_{k=1}^{p-1} \eta_i \eta_k \left(\sum_{t=1}^{n}\sum_{s=1}^{t-1} \cos^2\left((t-s)\lambda_j\right) \right) \frac{1}{\pi n m} g_{ik} c \frac{\lambda_m}{1 - d_i - d_k - 2d_e}$$

$$= \quad \sum_{i=1}^{p-1}\sum_{k=1}^{p-1} \eta_i \eta_k \frac{n}{4\pi m} g_{ik} c \frac{\lambda_m}{1 - d_i - d_k - 2d_e},$$

since $\sum_{t=1}^{n-1}\sum_{s=1}^{n-t} \cos^2\left(s\lambda_j\right) = (n-1)^2/4$. Thus, we finally arrive at

$$(35) = \sum_{i=1}^{p-1}\sum_{k=1}^{p-1} \eta_i \eta_k \frac{1}{2} g_{ik} c \frac{1}{1 - d_i - d_k - 2d_e} = \sum_{i=1}^{p-1}\sum_{k=1}^{p-1} \eta_i \eta_k c J_{ik},$$

where

$$J_{ik} = \frac{g_{ik}}{2\left(1 - d_i - d_k - 2d_e\right)}, \quad 1 \leq i, k \leq p-1, \tag{36}$$

and we have shown (27).

To show (29),

$$\sum_{t=1}^{n} E\left(z_{tn}^4\right) = \sum_{t=1}^{n} E\left(\sum_{s=1}^{t-1} \varepsilon_s' c_{t-s,n} \varepsilon_t \varepsilon_t' \sum_{r=1}^{t-1} c_{t-r,n}\varepsilon_r \sum_{p=1}^{t-1} \varepsilon_p' c_{t-p,n}\varepsilon_t\varepsilon_t' \sum_{q=1}^{t-1} c_{t-q,n}\varepsilon_q \right)$$

$$\leq \quad C\left(\sum_{t=1}^{n} \mathrm{tr}\left(\sum_{s=1}^{t-1} c_{t-s,n}' c_{t-s,n} c_{t-s,n}' c_{t-s,n} \right) + \sum_{t=1}^{n} \mathrm{tr}\left(\sum_{s=1}^{t-1} c_{t-s,n}' \sum_{r=1}^{t-1} c_{t-r,n} c_{t-r,n}' c_{t-s,n} \right) \right),$$

for some constant $C > 0$, by Assumption B'. This expression can be bounded by $O\left(n\left(\sum_{t=1}^{n} \left\| c_{tn}^2 \right\|\right)^2 \right) = O\left(n^{-1}\right)$, which completes the proof of asymptotic normality of $\hat{F}_{xe}$.

Hence, we have shown that

$$\sqrt{m}\lambda_m^{d_e-1} \tilde{\Lambda}^{-1} \hat{F}_{xe} \xrightarrow{d} N\left(0, cJ\right) \tag{37}$$

and the theorem follows using (18), (19), and (37) and applying the Slutsky Theorem. ∎

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.

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

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

Bandi, F. M. & Perron, B. (2002), 'A fundamentally different interpretation of the relation between implied and realized volatility', *Preprint, Universite de Montreal* .

Beran, J. (1994), *Statistics for Long-Memory Processes*, Chapman-Hall.

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637–654.

Brillinger, D. R. (1981), *Time Series, Data Analysis and Theory*, Holden Day, Inc., San Francisco.

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

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

Dahlhaus, R. (1989), 'Efficient parameter estimation for self-similar processes', *Annals of Statistics* **17**, 1749–1766.

Dolado, J. & Marmol, F. (1996), 'Efficient estimation of cointegrating relationships among higher order and fractionally integrated processes', *Working Paper no. 9617, Banco de Espana* .

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

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

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

Geweke, J. & 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.

Granger, C. W. J. & Joyeux, R. (1980), 'An introduction to long memory time series models and fractional differencing', *Journal of Time Series Analysis* **1**, 15–39.

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

Hosking, J. R. M. (1981), 'Fractional differencing', *Biometrika* **68**, 165–176.

Hull, J. C. (1997), *Options, Futures, and Other Derivatives*, third edn, Prentice-Hall, Inc.

Künsch, H. R. (1987), Statistical aspects of self-similar processes, *in* Y. Prokhorov & V. V. Sazanov, eds, 'Proceedings of the First World Congress of the Bernoulli Society', VNU Science Press, Utrecht, pp. 67–74.

Lobato, I. N. (1997), 'Consistency of the averaged cross-periodogram in long memory series', *Journal of Time Series Analysis* **18**, 137–155.

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

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

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

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

Merton, R. C. (1973), 'Theory of rational option pricing', *Bell Journal of Economics and Management Science* **4**, 141–183.

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

Rubinstein, M. & Vijh, A. M. (1987), 'The Berkeley options data base: A tool for empirical research', *Advances in Futures and Options Research* **2**, 209–221.

Shimotsu, K. & Phillips, P. C. B. (2002), 'Exact local whittle estimation of fractional integration', *Cowles Foundation Discussion Paper 1367* .

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

Table 1: Summary Statistics

	Realized volatility $y_t = \ln \sigma_{RV,t}$	Implied volatility $x_t = \ln \sigma_{IV,t}$
Mean	-2.642928	-2.035616
Variance	0.122568	0.200101
Skewness	0.585426	-0.916736
Excess Kurtosis	0.457148	2.807389

This table reports summary statistics for log realized and log implied volatility.

Table 2: Fractional Integration Order

	Realized volatility $y_t = \ln \sigma_{RV,t}$	Implied volatility $x_t = \ln \sigma_{IV,t}$
Box-Pierce (lags)[a]		
$Q\ (L=4)$	394.7^{**}	136.8^{**}
$Q\ (L=24)$	1211^{**}	649.4^{**}
GSP (bandwidth)[b]		
$\hat{d}\ (m=20)$	0.48468 (0.11180)	0.46281 (0.11180)
$\hat{d}\ (m=37)$	0.44758 (0.08220)	0.45273 (0.08220)
$\hat{d}\ (m=68)$	0.41620 (0.06063)	0.35033 (0.06063)
ADF (lags)[c]		
$t\ (L=0)$	-11.079^{**}	-14.843^{**}
$t\ (L=4)$	-4.6995^{**}	-5.3166^{**}

[a]Box-Pierce tests of the significance of the autocorrelation function. Two asterisks indicate significance at the 1% level in the asymptotic $\chi^2 (L)$ distribution.

[b]GSP are Gaussian semiparametric estimates of the fractional integration order as described in Robinson (1995a). Numbers in parentheses are asymptotic standard errors using $\sqrt{m}(\hat{d} - d) \to_d N (0, 1/4)$.

[c]ADF are Augmented Dickey-Fuller tests of the null of a unit root. Two asterisks indicate significance at the 1% level where the critical value is -3.44 (constant term included).

Table 3: Fractional Cointegration Analysis

Bandwidth	$\hat{\alpha}_m$	$\hat{\beta}_m$	$s.e.(\hat{\beta}_m)$			$\hat{d}_e$		
			$m_e = 20$	$m_e = 37$	$m_e = 68$	$m_e = 20$	$m_e = 37$	$m_e = 68$
$m = n - 1$	-1.8628	0.38325	$-$	$-$	$-$	0.38996 (0.11180)	0.33443 (0.08270)	0.31102 (0.06063)
$m = 15$	-0.91075	0.85094	0.33816	0.21970	0.18609	0.10991 (0.11180)	0.10048 (0.08270)	0.09479 (0.06063)
$m = 9$	-0.93154	0.84072	0.59762	0.26944	0.22107	0.11400 (0.11180)	0.10330 (0.08270)	0.09752 (0.06063)
$m = 6$	-0.93049	0.84124	0.63190	0.29297	0.24026	0.11378 (0.11180)	0.10315 (0.08270)	0.09738 (0.06063)
$m = 3$	-0.83362	0.88882	0.29858	0.25772	0.22412	0.09856 (0.11180)	0.09284 (0.08270)	0.08669 (0.06063)

This table reports semiparametric estimates of the stationary fractional cointegration relation (17) for different bandwidths. Bandwidth $m = n - 1$ corresponds to OLS. The standard error for $\hat{\beta}_m$ is based on the new asymptotic distribution result Theorem 2. The estimated fractional integration orders of the residuals $\hat{d}_e$ are GSP estimates with bandwidth $m_e = 20$, $m_e = 37$ and $m_e = 68$, respectively. For the integration orders of the raw series (used in the calculation of standard errors of $\hat{\beta}_m$), the results for bandwidth 68 from Table 2 were used.

Figure 1: Autocorrelation function up to lag 100 for log realized volatility
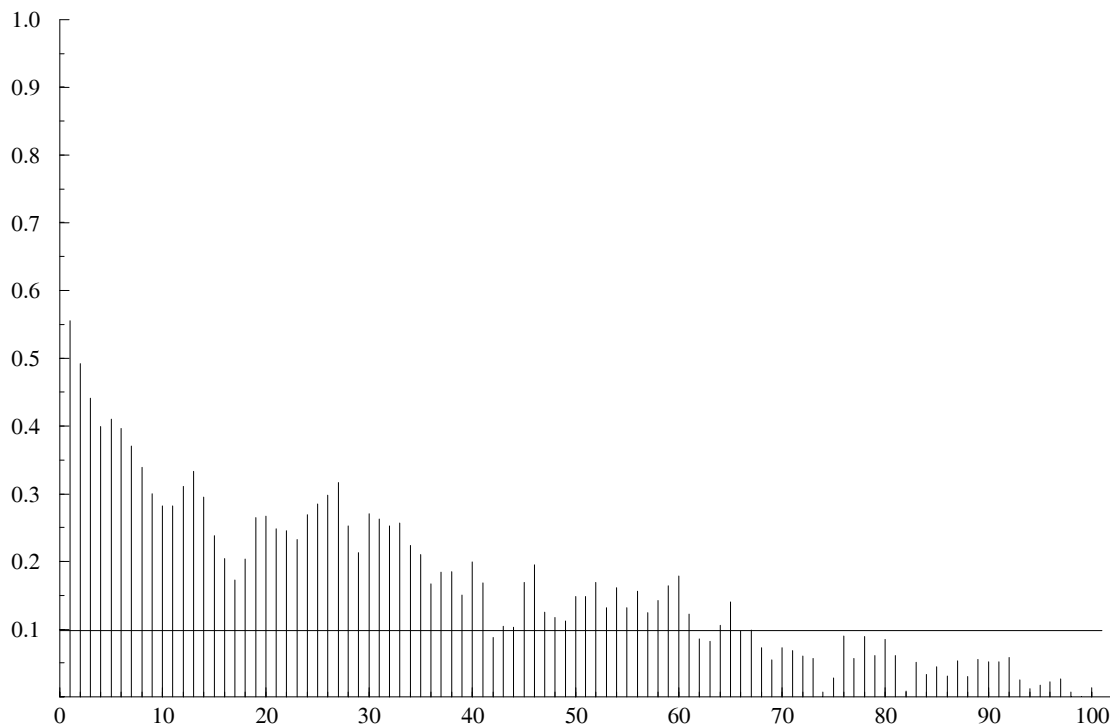

29

Figure 2: Autocorrelation function up to lag 100 for log implied volatility
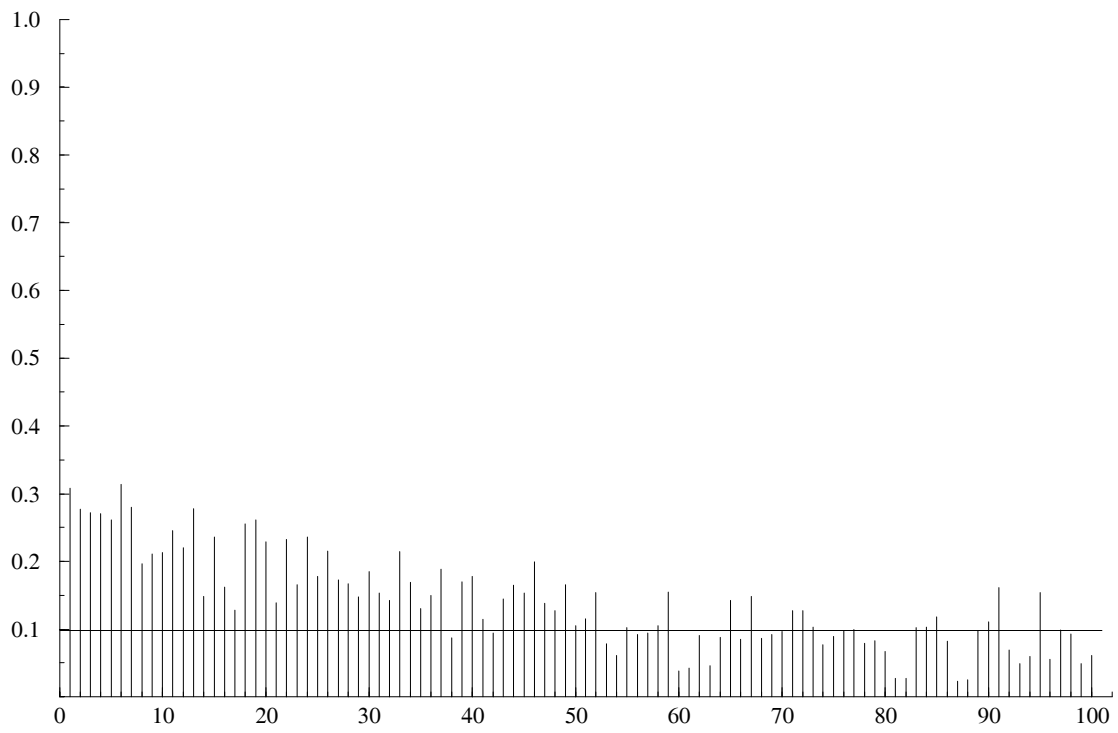
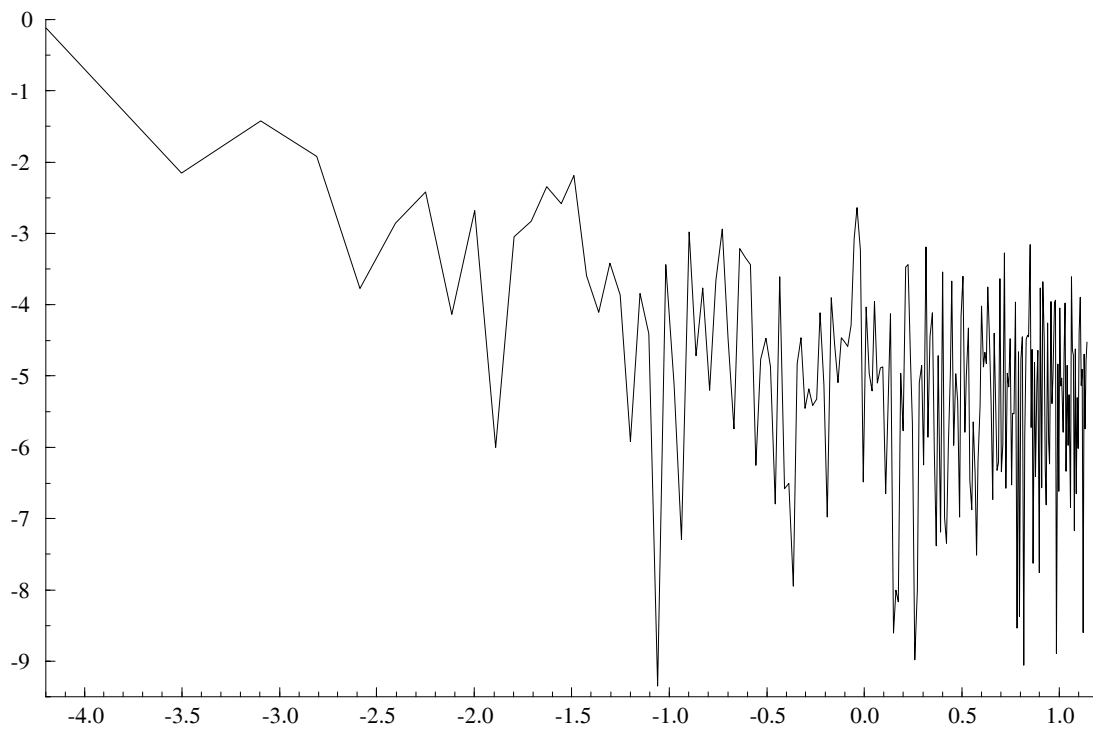Figure 3: Log periodogram vs log Fourier frequencies for log realized volatility

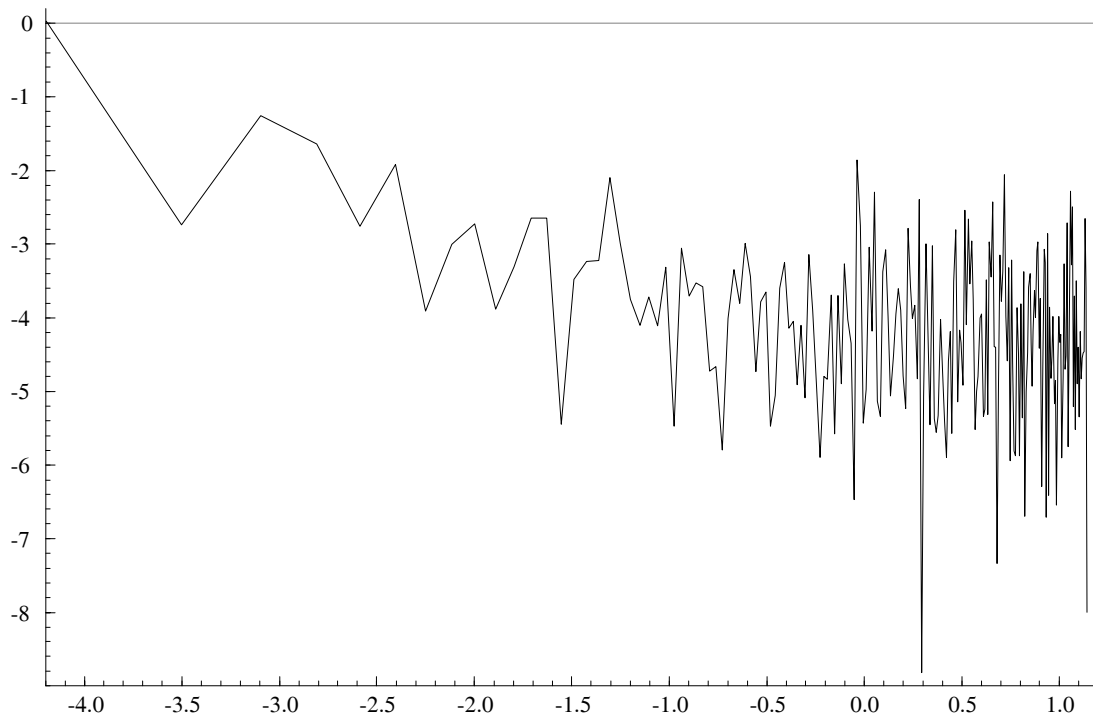Figure 4: Log periodogram vs log Fourier frequencies for log implied volatility

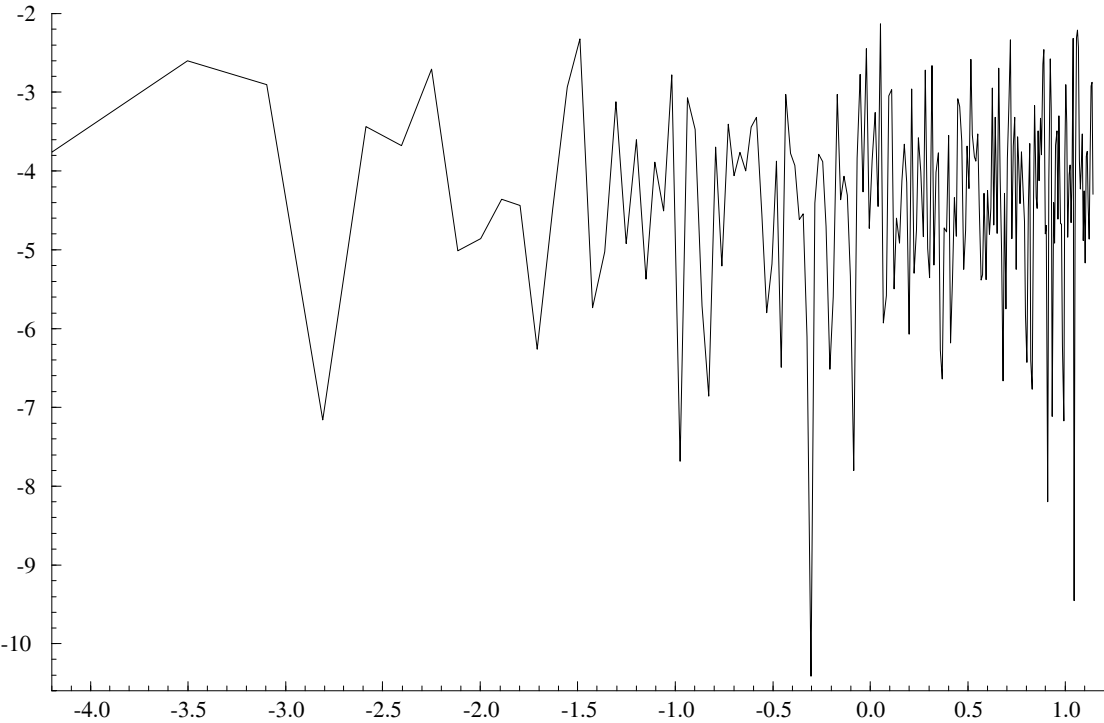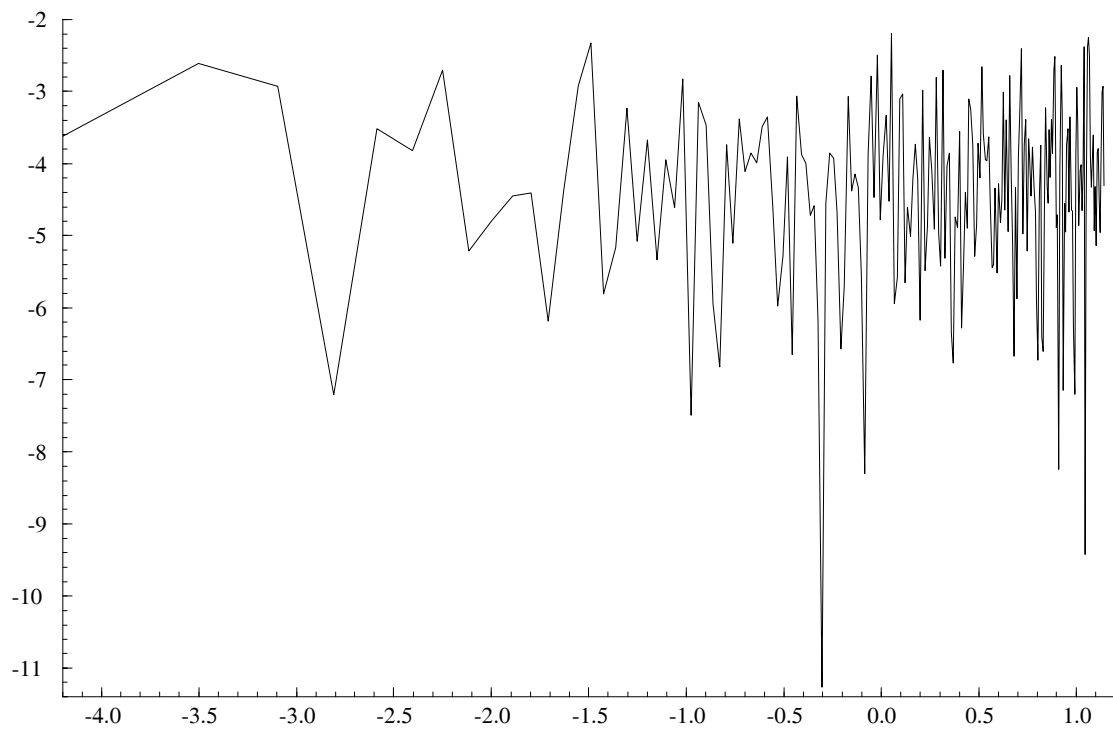Figure 5: Log periodogram vs log Fourier frequencies for residuals with m=3

Figure 6: Log periodogram vs log Fourier frequencies for residuals with m=15

CENTRE FOR DYNAMIC MODELLING IN ECONOMICS

DEPARTMENT OF ECONOMICS - UNIVERSITY OF AARHUS - DK - 8000 AARHUS C - DENMARK
☎ +45 89 42 11 33 - TELEFAX +45 86 13 63 34

Working papers, issued by the Centre for Dynamic Modelling in Economics:

1999-21   Martin Paldam: Corruption and Religion. Adding to the Economic Model?

2000-1   Niels Haldrup and Peter Lildholdt: On the Robustness of Unit Root Tests in the Presence of Double Unit Roots.

2000-2   Niels Haldrup and Peter Lildholdt: Local Power Functions of Tests for Double Unit Roots.

2000-3   Jamsheed Shorish: Quasi-Static Macroeconomic Systems.

2000-4   Licun Xue: A Notion of Consistent Rationalizability - Between Weak and Pearce's Extensive Form Rationalizability.

2000-6   Graham Elliott and Michael Jansson: Testing for Unit Roots with Stationary Covariates.

2000-8   Niels Haldrup, Antonio Montanés and Andreu Sanso: Measurement Errors and Outliers in Seasonal Unit Root Testing.

2000-12   Effrosyni Diamantoudi and Licun Xue: Farsighted Stability in Hedonic Games.

2000-13   Licun Xue: Stable Agreements in Infinitely Repeated Games.

2000-16   Boriss Siliverstovs: The Bi-parameter Smooth Transition AutoRegressive model.

2000-17   Peter Skott: Demand Policy in the Long Run.

2000-18   Paul Auerbach and Peter Skott: Skill Asymmetries, Increasing Wage Inequality and Unemployment.

2000-19   Torben M. Andersen: Nominal Rigidities and the Optimal Rate of Inflation.

2001-1   Jakob Roland Munch and Michael Svarer: Mortality and Socio-economic Differences in a Competing Risks Model.

2001-4   Bent Jesper Christensen and Morten Ø. Nielsen: Semiparametric Analysis of Stationary Fractional Cointegration and the Implied-Realized Volatility Relation in High-Frequency Options Data.

Working Paper

2000-11   Michael Rosholm and Michael Svarer: Structurally Dependent Competing Risks.

2000-12   Efforsyni Diamantoudi and Licun Xue: Farsighted Stability in Hedonic Games.

2000-13   Licun Xue: Stable Agreements in Infinitely Repeated Games.

2000-14   Bo Sandemann Rasmussen: Government Debt and Capital Accumulation in the Blanchard-Cass-Yaari OLG Model.

2000-15   Nikolaj Malchow-Møller and Bo Jellesmark Thorsen: Investment under Uncertainty - the Case of Repeated Investment Options.

2000-16   Boriss Siliverstovs: The Bi-parameter Smooth Transition AutoRegressive model.

2000-17   Peter Skott: Demand Policy in the Long Run.

2000-18   Paul Auerbach and Peter Skott: Skill Assymmetries, Increasing Wage Inequality and Unemployment.

2000-19   Torben M. Andersen: Nominal Rigidities and the Optimal Rate of Inflation.

2001-1   Jakob Roland Munch and Michael Svarer: Mortality and Socio-economic Differences in a Competing Risks Model.

2001-2   Effrosyni Diamantoudi: Equilibrium Binding Agreements under Diverse Behavioral Assumptions.

2001-3   Bo Sandemann Rasmussen: Partial vs. Global Coordination of Capital Income Tax Policies.

200-4   Bent Jesper Christensen and Morten Ø. Nielsen: Semiparametric Analysis of Stationary Fractional Cointegration and the Implied-Realized Volatility Relation.