Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting

B.J. Christensen and M. Ørregaard Nielsen
Asymptotic Normality of Narrow-Band Least Squares in the Stationary Fractional Cointegration Model and Volatility Forecasting

Bent Jesper Christensen
University of Aarhus
E-mail: bjchristensen@econ.au.dk

Morten Ørregaard Nielsen*
Cornell University
E-mail: mon2@cornell.edu

April 5, 2004

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 is due to Robinson (1994, *Annals of Statistics* 22, 515-539) and 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. We derive the asymptotic distribution theory for the FDLS estimator of the cointegration vector in the stationary long memory case, thus complementing Robinson’s consistency result. An application to the relation between the volatility realized in the stock market and the associated implicit volatility derived from option prices is offered.

*JEL classification:* C14; C22; G13

*Keywords:* Asymptotic distribution theory; High-frequency data; Long memory; Semiparametric methods; Stationary fractional cointegration

*Corresponding author. Please address correspondence to: Morten Ørregaard Nielsen, Department of Economics, Cornell University, 482 Uris Hall, Ithaca, NY 14853, USA; phone: +1 607 255 6338; fax: +1 607 255 2818; email: mon2@cornell.edu.
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 $\beta$ such that

$$ y_t = \beta' x_t + e_t, $$

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

$$ (1 - L)^d x_t = \varepsilon_t, $$

1Note that (2) is valid only for $d < 1/2$. For nonstationary values of $d$, $I(d)$ is defined either by initialization or by considering partial sums of series of the type in (2), see e.g. Marinucci & Robinson (1999).
where $\varepsilon_t \in I(0)$ and $(1 - L)^d$ is defined by its binomial expansion

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

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), Dueker & Startz (1998), and Robinson & Hualde (2003), among others, using parametric methods. However, Robinson (1994b) shows that conventional estimators, and in particular OLS, are inconsistent when the errors are fractionally integrated. He introduces narrow-band least squares, a semiparametric method, and proves that it is consistent even in situations where the error term is correlated with the regressors, e.g., as a result of fractional cointegration. Robinson’s method has been extended to multiple regressors by Lobato (1997) and further developed by Robinson & Marinucci (2003) (henceforth RM) and applied by Marinucci & Robinson (2001). These 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).

The distribution theory for fractional cointegration is not complete for all relevant parameter regions. 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 provide the limiting distribution for the case $d > 0, d_e \geq 0, d + d_e < 1/2$, which has been missing in the literature. This is the case of stationary fractional cointegration. Precisely in this case, Robinson’s result yields inconsistency of conventional estimators, including OLS. Thus, in the stationary fractional cointegration
model narrow-band least squares is a tool for consistent estimation of the regression parameters even in situations when the error term is correlated with the regressors.

The present paper complements the consistency result for narrow-band least squares of Robinson (1994b) and the distribution results of RM by providing the relevant distribution theory for the stationary fractional cointegration model. 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. Fox & Taqqu (1986, 1987) and Beran (1994).

We illustrate our new distribution theory in an application to stock market 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). For univariate stock market volatility, Andersen, Bollerslev, Diebold & Ebens (2001) show that the parameter region \( 0 < d < 1/2 \) is relevant. We consider the multivariate case, using volatility series measured from both stock index returns (realized volatility) and options data (implied volatility). We examine the possibility that the two volatility series are fractionally cointegrated, consistent with the notion that implied volatility forecasts realized volatility. Our analysis shows 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. The condition \( d + d_e < 1/2 \) implies that the limiting distributions provided by RM do not apply. Using our new distribution theory, we find evidence in favor of long-run unbiasedness of the implied volatility forecast of subsequent realized volatility.

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 \( \{x_t, t = 1, 2, \ldots\} \) generated by (2) has spectral density

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

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 long memory since the autocorrelations die out at a hyperbolic rate, in contrast to the much faster exponential rate in, e.g., the stationary and invertible ARMA\((p,q)\) case. The most well-known model satisfying (4) is the fractional ARIMA model of Granger & Joyeux (1980) and Hosking (1981). For excellent surveys of long memory processes and fractional models see Robinson (1994c), Baillie (1996), and Robinson (2003), and for a textbook treatment 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,
1987) 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

$$y_t = \beta' x_t + e_t.$$  \hspace{1cm} (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 (1994b) and subsequently Robinson & Marinucci (2003) and Marinucci & Robinson (2001) attempt to correct this by considering a semiparametric approach to the estimation of the cointegration equation (5). Other recent references for fractional cointegration analyses include *inter alia* Chen & Hurvich (2003a, 2003b) and Velasco (2003), who also consider semiparametric approaches. However, they all consider only 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\(^2\). 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 (2003) 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 \( z_t = (x'_t, y_t)' \), \( t = 1, \ldots, n \), which satisfy the following assumptions.

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

\[
    f_{zz}(\lambda) \sim AGA \quad \text{as} \quad \lambda \to 0^+,
\]

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

\[
    \Lambda = diag \{ \lambda^{-d_1}, \ldots, \lambda^{-d_p} \}
\]

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

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

\(^2\)Cointegration in the stationary case is also considered by Marinucci (2000), where a continuously averaged (integral) version of FDLS, see (7) below, is advocated. However, only consistency is shown in the stationary case and no asymptotic distribution theory is given. Furthermore, the approach of the present paper does not require mean correction and is therefore probably superior.
with \( 0 \leq d_e < \min (d_i) \).

**Assumption B:** \( z_t \) admits the Wold representation

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

where the innovations \( \varepsilon_t \) satisfy

\[
E(\varepsilon_t | \mathcal{F}_{t-1}) = 0, \quad E(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}) = \Sigma, \ a.s.,
\]

and \( \varepsilon_t \varepsilon_t' \) are uniformly integrable. Here, \( \mu_z = E(z_0) \), \( \Sigma \) is a constant matrix of full rank, \( \mathcal{F}_t = \sigma(\{\varepsilon_s, s \leq t\}) \) is the \( \sigma \)-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 an 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

\[
I_{ab}(\lambda) = w_a(\lambda) w_b^*(\lambda) = I_{ab}^c(\lambda) + iI_{ab}^q(\lambda),
\]

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

\[
\hat{F}_{ab}(k,l) = \frac{2\pi}{n} \sum_{j=k}^{l} I_{ab}^c(\lambda_j), \quad 1 \leq k \leq l \leq n - 1,
\]  \( \text{(6)} \)
for $\lambda_j = 2\pi j/n$. If $\hat{F}_{ab}$ is a vector we denote the $i$’th entry $\hat{F}^{(i)}_{ab}$, and if it is a matrix we denote the $(i, j)$’th entry $\hat{F}^{(i,j)}_{ab}$. We could also have considered a continuously averaged version of (6) as in Marinucci (2000), but it would not enjoy the property of being invariant to mean terms.

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

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

of $\beta$ in the regression (5). Notice that by this definition $\hat{\beta}_{n-1}$ is algebraically identical to the usual OLS estimate of $\beta$ with allowance for a non-zero mean in $e_t$. If

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

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.

In this setup, RM show the following result.

**Theorem 1 (Robinson and Marinucci, 2003)** Under Assumptions A, B, (8), and with $\hat{\beta}_m$ defined in (7),

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

Notice that under fractional cointegration $d_e < \min(d_i)$, so the estimator $\hat{\beta}_m$ is consistent for $\beta$. Furthermore, if the integration order of the raw data series is common, i.e. $d_i = d$ for all $i = 1, \ldots, p$, the stochastic order of magnitude 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, \( f(\lambda) = \{ f_{ij}(\lambda) \}_{i,j=1,...,p} \) 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

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

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

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

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(\varepsilon_t|\mathcal{F}_{t-1}) = 0, \ E(\varepsilon_t\varepsilon'_t|\mathcal{F}_{t-1}) = I_p, \ \text{and the matrices} \ E(\varepsilon_t \otimes \varepsilon_t'|\mathcal{F}_{t-1}), \ E(\varepsilon_t\varepsilon'_t \otimes \varepsilon_t\varepsilon'_t|\mathcal{F}_{t-1}) \ \text{are nonstochastic, finite, and do not depend on} \ t, \ \text{with} \ \mathcal{F}_t = \sigma(\{\varepsilon_s, s \leq t\}). \ Denote the periodogram of \( \varepsilon_t \) by \( J(\lambda) \).

**Assumption C**: As \( \lambda \to 0^+ \)

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

for \( i = 1, ..., p \), where \( A_i(\lambda) \) is the \( i'\text{th} \) 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 \( \alpha \) 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, 1987) or Dahlhaus (1989). Assumption A’ is satisfied with \( \alpha = 2 \) if, for instance, \( w_t \) is a vector fractional ARIMA process. Here, (10) with \( g_{ip} = g_{pi} = 0, i = 1, \ldots, p - 1 \), is a central assumption, restricting the coherencies between the cointegration residuals and other right hand side variables at the origin. The restriction is not as strong as it may seem at first, and in particular it does not require the regressors \( x_t \) and the errors \( e_t \) to be uncorrelated at frequencies away from the origin. Thus, our estimator allows the regressor and error terms to share the same short- and medium-term dynamics. 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 \( \beta \) 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 \( e_t \) form an i.i.d. process with finite fourth moments. Under Assumption B’ we can write the spectral density matrix of \( w_t \) as

\[
f(\lambda) = \frac{1}{2\pi} A(\lambda) A^*(\lambda),
\]

(11)

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 \( \alpha \) is high, i.e. (9) is a close approximation to (11), 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, \( d_i + d_e < 1/2 \), \( 0 \leq d_e < d_i < 1/2 \) for \( i = 1, \ldots, p - 1 \), and with \( \hat{\beta}_m \) defined in (7),

\[
\sqrt{m\lambda_m^{d_e}} \tilde{\Lambda}_m(\hat{\beta}_m - \beta) \xrightarrow{d} N \left( 0, g_{pp}H^{-1}JH^{-1} \right)
\]
as $n \to \infty$, where $\hat{\Lambda}_m = \text{diag}(\lambda_m^{-d_1}, \ldots, \lambda_m^{-d_{p-1}})$ and $H, J$ are defined in equations (21) and (37) in the appendix.

Note that the limiting distribution in Theorem 2 is a centered normal, unlike the non-normal distributions arising when $d_i \geq 1/2$ (Marinucci & Robinson (2001) and Robinson & Marinucci (2003)). The particular result that the asymptotic bias of $\hat{\beta}_m$ disappears is due to the condition (10) and $g_{ip} = g_{pi} = 0$, $i = 1, \ldots, p - 1$, in Assumption A'. The fact that the limiting distribution is normal is to some extent expected, in view of the results from Fox & Taqqu (1986, 1987) and subsequent authors, showing that quadratic forms of long memory processes with square-summable autocovariances ($2d < 1/2$) are asymptotically Gaussian. This is related to our result, where we work with a quadratic form with $d_i + d_e < 1/2$. See also Lobato & Robinson (1996). Our asymptotic normality result allows for straightforward inference, e.g. based on standard $t$-ratios (no simulation needed).

In the simple two-variable case with $x_t \in I(d_x)$, $e_t \in I(d_e)$, and $G = \text{diag}(g_x, g_e)$, the asymptotic distribution of $\hat{\beta}_m$ is given by Theorem 2 as

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

The asymptotic distribution (12) is easy to interpret and compare to well-known cases. The parameters $g_e$ and $g_x$ correspond to (long-run versions of) $\text{var} (\Delta^{d_e} e_t)$ and $\text{var} (\Delta^{d_x} x_t)$. Ordinary full band spectral regression yields the asymptotic variance $g_e/g_x$ 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 (12). This arises since the cross-terms involving $e_t$ vanish asymptotically in $\text{var}(\hat{F}_{xe})$ (see the proof of Theorem 2, immediately before (36)). 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 $d_x = d_e = 0$ of (12) and Theorem 2.

The parameters $G, d_e, d_i, ~i = 1, \ldots, p - 1$, 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 as-
ymptotic 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 \( \hat{b} = \hat{d} - \hat{d}_e \) (assuming a common integration order \( d_i = d \) for the raw data series) using e.g. the methods mentioned above. Formally, Velasco (2003) 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 \( i = 1, ..., p - 1 \) 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 distribution like the scalar averaged periodogram estimator in Lobato & Robinson (1996).

To sum up, Theorem 2 provides the necessary asymptotic distribution theory for the stationary fractional cointegration model, thus complementing the consistency result of Robinson (1994b) and Robinson & Marinucci (2003). We use this in the empirical application below.

3 The Implied-Realized Volatility Relation

Our application of stationary fractional cointegration analysis is to the relation between the volatility implied in option prices and the subsequently 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 expected 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 $\tau$ periods to expiration and strike price $k$ is given by

\[ c(s,k,\tau,r,\sigma) = s\Phi(\delta) - e^{-\tau r}k\Phi(\delta - \sigma\sqrt{\tau}), \quad (13) \]

\[ \delta = \frac{\ln (s/k) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}, \]

where $s$ is the price of the underlying asset, $r$ is the riskless interest rate, $\Phi$ is the standard normal c.d.f., and $\sigma$ is the return volatility of the underlying asset through expiration of the option ($\tau$ periods hence).

Given an observation $c$ of the price of the option, the implied volatility $\sigma_{IV}$ may be determined by inverting (13), i.e. solving the nonlinear equation

\[ c = c(s,k,\tau,r,\sigma_{IV}) \quad (14) \]

numerically with respect to $\sigma_{IV}$, for given data on $s,k,\tau,$ 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 $t + \tau$. Realized volatility is simply computed as the sample standard deviation $\sigma_{RV,t}$ of the realized return from $t$ to $t + \tau$.

Christensen & Prabhala (1998) considered the regression specification

\[ y_t = \alpha + \beta x_t + e_t, \quad (15) \]

where $y_t = \ln \sigma_{RV,t}$ and $x_t = \ln \sigma_{IV,t}$ are the log-volatilities, and the unbiasedness hypothesis of interest is that $\beta = 1$. 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. Basic OLS regression in (15) produced a $\beta$-estimate less than unity. Christensen & Prabhala (1998) also presented results without the log transform, and the difference was negligible.

If volatility is persistent and, indeed, fractionally integrated, as empirical literature suggests (Andersen, Bollerslev, Diebold & Ebens (2001) find fractional integration with $d$ at about $0.35 - 0.4$), whereas the forecasting error $e_t$ in (15) is serially uncorrelated or possesses only short memory, then $x_t$ and $y_t$ are fractionally cointegrated. The condition $\beta = 1$ is then a long-run unbiasedness hypothesis. However, from Robinson (1994b) and Robinson & Marinucci...
OLS is generally inconsistent for $\beta$ in the case of stationary fractional cointegration. The narrow-band analysis is consistent and the results from the previous section provide the necessary asymptotic distribution theory for inference on $\beta$.

Our empirical results below in fact support fractional cointegration of $x_t$ and $y_t$. This is consistent with Bandi & Perron (2002), who apply narrow band analysis to the implied-realized volatility relation in independent contemporaneous work. Due to the lack of available asymptotic distribution theory, they rely on subsampling for inference, and find results consistent with $\beta = 1$. Our inference below is based on asymptotic standard errors from the new theory (Theorem 2), and also agrees with this long-run unbiasedness hypothesis.

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 (13), the SPX options have the advantage over the OEX contracts considered by Christensen & Prabhala (1998) 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 Christensen & Prabhala (1998) 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

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

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. For increasing sampling frequency, $K\sigma^2_{RV,t}$ converges to the true integrated variance, see Andersen, Bollerslev, Diebold & Ebens (2001) and Barndorff-Nielsen & Shephard (2002).

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 (13), 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
trading days) to expiration then $\tilde{\sigma}_{IV,t-i}$ corresponds to $i + 1$ weeks ($\kappa_i = 5i + 4\frac{5}{6}$ 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 $\kappa_0$) intervals by the formula

$$\sigma_{IV,t-i}^2 = \frac{1}{\kappa_i - \kappa_{i-1}} \left( \kappa_i \cdot \tilde{\sigma}_{IV,t-i}^2 - \kappa_{i-1} \cdot \tilde{\sigma}_{IV,t-i+1}^2 \right),$$

(17)

starting with $\sigma_{IV,t} = \tilde{\sigma}_{IV,t}$ for $t$ corresponding to a one week option and applying (17) for $i = 1, 2, 3$ and if applicable also $i = 4$. Here, (17) is an identity for the associated realized volatilities covering interval lengths $\kappa_0$, $\kappa_i$, and $\kappa_{i-1}$. This identity becomes an approximation for implied (as opposed to realized) volatilities. In practice, if the RHS of (17) 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. By focusing on a diminishing band of low frequencies in the fractional cointegration analysis, the effect of such measurement errors vanishes asymptotically, provided the spectral density of these errors is dominated by that of the observed variables, i.e. is of lower order of fractional integration.

4.2 Empirical Results

Summary statistics for the two (log) 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 (13) 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 (13)).

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 (17), 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
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**

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

$$y_t = \alpha + \beta x_t + e_t$$

(18)
of realized on implied volatility. The parameter of interest, $\beta$, 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 stationary fractional cointegration is recognized, but the narrow-band estimator remains consistent even in this case (Robinson, 1994).

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, $\beta$, we find much larger point estimates. Following Marinucci & Robinson (2001) and Robinson & Marinucci (2003), 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, bandwidth 68, for $d$. We have chosen bandwidth 68 since the results are quite close and this choice yields results closest to those in the literature. 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, which is also in accordance with Bandi & Perron (2002). The residuals are of lower order of fractional integration than the volatility

---

3The rationale for this has been provided by Robinson & Marinucci (2001) and Chen & Hurvich (2003a, 2003b), who showed that a small (finite) $m$ is sufficient for consistency and produces much less bias. However, those results refer to the nonstationary case, where much more caution is needed. Thus, we consider also wider bandwidths up to $m = 15$, see also Robinson (1994a).
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**

For the estimation of $d_e$, 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 choice of bandwidth for the residuals. For example, the plots using $m = 3$ and $m = 15$ (to estimate $\beta$ 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 $\beta$, 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, based on the asymptotic theory from Theorem 2, which contains unity. 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. These findings based on the new asymptotic theory agree with those obtained by Bandi & Perron (2002) using subsampling, and are obviously in stark contrast to the raw OLS results in the first line of the table. Clearly, inferences on volatility relations may be heavily misleading if the possibility of stationary fractional cointegration is ignored.

5 Conclusion

In this paper, we have derived the asymptotic distribution theory for the narrow-band least squares estimator of Robinson (1994b) in the stationary fractional cointegration model. This
new distribution theory complements the consistency result for the cointegration coefficients from Robinson (1994b). The method is particularly useful because it provides consistent and asymptotically normal (CAN) estimators even in situations where OLS is inconsistent, which it is in the stationary fractional cointegration model due to the correlation between regressors and error terms. 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 an empirical application, we show that inference based on the new asymptotic theory supports long-run unbiasedness in the relation between implied and realized volatility, consistent with Bandi & Perron’s (2002) findings based on subsampling.

Acknowledgements

This paper is a revised version of our earlier paper “Semiparametric Analysis of Stationary Fractional Cointegration and the Implied-Realized Volatility Relation in High-Frequency Options Data”, August 2001, Department of Management Working Paper 2001-9, University of Aarhus, see http://www.econ.au.dk/afv/wp/wp-liste.htm. We thank participants at the Danish Econometric Society annual meeting, Denmark, 26-28 May 2000, the ‘Market Microstructure and High-Frequency Data in Finance’ conference, Denmark, 10-12 August 2001, and the ESRC Econometric Study Group Annual Conference (invited presentation), Bristol, 10-12 July 2003, for useful comments. We also thank Tim Bollerslev, two anonymous referees, and the associate editor for helpful suggestions. Finally, we are grateful to the Centre for Analytical Finance (CAF), the Centre for Dynamic Modelling in Economics (CDME), and the Danish Social Science Research Council (SSF) for research support.

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 $\varepsilon_t$ 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

\[ \sqrt{m \lambda_m} \tilde{\Lambda}_m (\tilde{\beta}_m - \beta) = \sqrt{m \lambda_m^d \tilde{\Lambda}_m} \tilde{F}_{xx}^{-1} (1, m) \tilde{F}_{xe} (1, m) \]

\[ = \left\{ \lambda_m \tilde{\Lambda}_m \tilde{F}_{xx}^{-1} (1, m) \tilde{\Lambda}_m \right\} \left\{ \sqrt{m \lambda_m^{d-1} \tilde{\Lambda}_m^{-1} \tilde{F}_{xe} (1, m)} \right\}. \]  

(19)

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

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

Thus, the first term in (19) is

\[ \lambda_m \tilde{\Lambda}_m \tilde{F}_{xx}^{-1} (1, m) \tilde{\Lambda}_m \overset{p}{\to} \lambda_m \tilde{\Lambda}_m \tilde{\Lambda}_m^{-1} \lambda_m^{-1} H^{-1} \tilde{\Lambda}_m^{-1} \tilde{\Lambda}_m = H^{-1} \]

(20)

by the Continuous Mapping Theorem, where

\[ H_{ik} = \frac{g_{ik}}{1 - d_i - d_k}, \ 1 \leq i, k \leq p - 1. \]

(21)

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

\[ \sqrt{m \lambda_m^{d-1} \tilde{\Lambda}_m^{-1} \tilde{F}_{xe} (1, m)}. \]

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

\[ \eta^t \sqrt{m \lambda_m^{d-1} \tilde{\Lambda}_m^{-1} \tilde{F}_{xe} (1, m)} = \sum_{i=1}^{p-1} \eta_i \sqrt{m \lambda_m^{d_i+d_e-1} \tilde{F}_{ip} (\lambda_m)} \]

\[ = \sum_{i=1}^{p-1} \eta_i \sqrt{m \lambda_m^{d_i+d_e-1}} \frac{12 \pi}{n} \sum_{j=1}^{m} \Re \left( \hat{I}_{ip} (\lambda_j) \right) \]

\[ = \sum_{i=1}^{p-1} \eta_i \sqrt{m \lambda_m^{d_i+d_e-1}} \frac{12 \pi}{n} \sum_{j=1}^{m} \Re \left( \hat{I}_{ip} (\lambda_j) - A_i (\lambda_j) J (\lambda_j) A_p^*(\lambda_j) \right) \]

\[ + \sum_{i=1}^{p-1} \eta_i \sqrt{m \lambda_m^{d_i+d_e-1}} \frac{12 \pi}{n} \sum_{j=1}^{m} \Re \left( A_i (\lambda_j) J (\lambda_j) A_p^*(\lambda_j) \right). \]

(22)

(23)
By (C.2) of Lobato (1999), which is implied by our assumptions, the first term, i.e. expression (22), is

\begin{align*}
(22) &= O_p \left( \sum_{i=1}^{p-1} \eta_i \sqrt{m \lambda_m^{-1}} \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.
\end{align*}

The second term above, expression (23), can be written

\begin{align*}
(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} \epsilon_t e^{it\lambda_j} 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} \epsilon_t e^{itA^*_p (\lambda_j)} \right) \\
&\quad + \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_{s \neq t}^{n} \epsilon_t \epsilon_s e^{i(t-s)\lambda_j} A^*_p (\lambda_j) \right) \quad (24)
\end{align*}

Now write (24) as

\begin{align*}
(24) &= \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) A^*_p (\lambda_j) \right) \quad (26) \\
&\quad + \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} \epsilon_t \epsilon_t' - I_p \right) A^*_p (\lambda_j) \right), \quad (27)
\end{align*}

where (26) is bounded by

\begin{align*}
\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^{\alpha-d_i-d_e}}{n} \right) \\
&= O \left( \frac{m^{1+2\alpha}}{n^{2\alpha}} \right) \to 0
\end{align*}

by Assumptions A’ and D. For (27) we use the fact that $\epsilon_t \epsilon_t' - I_p$ is a martingale difference sequence with respect to the filtration $(\mathcal{F}_t)_{t \in \mathbb{Z}}$, $\mathcal{F}_t = \sigma \left( \{ \epsilon_s, s \leq t \} \right)$, such that in particular
\[ n^{-1} \sum_{t=1}^{n} \varepsilon_t \varepsilon_t' - I_p = o_p(1). \]  Then (27) is bounded by
\[ \sup_t o_p \left( \sqrt{m} \lambda_m^{d_t + d_e - 1} \frac{1}{n} \sum_{j=1}^{m} f_{ip}(\lambda_j) \right) = o_p \left( \frac{m^{1+2\alpha}}{n^{2\alpha}} \right) \rightarrow_p 0 \]
by Assumptions A' and D.

We return to (25),
\[
\begin{align*}
\sum_{i=1}^{p-1} \eta_i \sqrt{m} \lambda_m^{d_t + d_e - 1} \frac{1}{n^2} \sum_{j=1}^{m} \text{Re} \left( A_i(\lambda_j) \sum_{t=1}^{n} \sum_{s \neq t} \varepsilon_t \varepsilon_s' e^{i(t-s)\lambda_j} A_p(\lambda_j) \right) \\
= \sum_{t=1}^{n} \varepsilon_t' \sum_{s \neq t} \eta_i \sqrt{m} \lambda_m^{d_t + d_e - 1} \sum_{j=1}^{m} \text{Re} \left( A_i'(\lambda_j) e^{i(t-s)\lambda_j} A_p(\lambda_j) \right) \varepsilon_s \\
= \sum_{t=1}^{n} \varepsilon_t' \sum_{s=1}^{t-1} c_{t-s,n} \varepsilon_s,
\end{align*}
\]
defining
\[
\begin{align*}
c_{tn} &= \frac{1}{2\pi n \sqrt{m}} \sum_{j=1}^{m} \theta_j \cos (t\lambda_j), \\
\theta_j &= \sum_{i=1}^{p-1} \eta_i \lambda_m^{d_t + d_e} \text{Re} \left( A_i'(\lambda_j) A_p(\lambda_j) + A'_p(\lambda_j) A_i(\lambda_j) \right).
\end{align*}
\]
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
\[
\begin{align*}
\sum_{t=1}^{n} E \left( z_{tn}^2 \mid \mathcal{F}_{t-1} \right) - \sum_{i=1}^{p-1} \sum_{k=1}^{p-1} \eta_i \eta_k g_{pp} J_{ik} \rightarrow_p 0, \quad (28) \\
\sum_{t=1}^{n} E \left( z_{tn}^2 1(|z_{tn}| > \delta) \right) \rightarrow 0, \quad \delta > 0, \quad (29)
\end{align*}
\]
see Hall & Heyde (1980, chap. 3.2). A sufficient condition for (29) is
\[
\sum_{t=1}^{n} E \left( z_{tn}^4 \right) \rightarrow 0. \quad (30)
\]
First, we show (28),

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

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

By (D.10) and (D.11) of Lobato (1999), (32) has mean zero and variance

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

Since \( \| \theta_j \| = O \left( 1 \right) \) by construction, \( \| c_{sn} \| \) is bounded by

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

Define the functions \( K_i(\lambda) = \text{Re} \left( A_i' (\lambda) \bar{A}_i (\lambda) + A_i' (\lambda) \bar{A}_i (\lambda) \right) \) such that \( \theta_j = \sum_{i=1}^{p-1} \eta_i \lambda^{d_i+d_e} K_i(\lambda_j) \), \( K_i(\lambda) = O \left( \lambda^{-d_i-d_e} \right) \) as \( \lambda \to 0^+ \), and \( K_i(\lambda) \) is differentiable with \( \partial K_i(\lambda) / \partial \lambda = O \left( \lambda^{-d_i-d_e-1} \right) \) as \( \lambda \to 0^+ \) by Assumption 3. Now we can derive an alternative bound as

\[
\| c_{sn} \| = O \left( \sup_{i} \frac{\lambda^{d_i+d_e}}{n \sqrt{m}} \sum_{j=1}^{m} K_i(\lambda_j) \cos \left( s \lambda_j \right) \right)
\]

\[
= O \left( \sup_{i} \frac{\lambda^{d_i+d_e}}{n \sqrt{m}} \sum_{j=1}^{m-1} (K_i(\lambda_j) - K_i(\lambda_{j+1})) \sum_{k=1}^{j} \cos \left( s \lambda_k \right) \right)
\]

\[
+ O \left( \sup_{i} \frac{\lambda^{m}}{n \sqrt{m}} K_i(\lambda_m) \sum_{j=1}^{m} \cos \left( s \lambda_j \right) \right)
\]

from summation by parts. Using the Mean Value Theorem it follows that \( K_i(\lambda_j) - K_i(\lambda_{j+1}) = \)

26
\[(\lambda_{j+1} - \lambda_j) \frac{\partial K_i(\lambda_j)}{\partial \lambda} = \frac{2\pi}{n} \frac{\partial K_i(\lambda_j)}{\partial \lambda}, \text{ and the bound is} \]

\[
||c_{sn}|| = O \left( \sup_i \frac{\lambda_m^{d_i+d_e}}{n \sqrt{m}} \sum_{j=1}^{m-1} \sum_{k=1}^{n-1} \lambda_j^{-d_i-d_e-1} \cos(s \lambda_k) \right) \\
+ O \left( \sup_i \frac{\lambda_m^{d_i+d_e}}{n \sqrt{m}} \lambda_m^{-d_i-d_e-1} \sum_{j=1}^{m} \cos(s \lambda_j) \right) \\
= O \left( \frac{1}{s \sqrt{m}} \right)
\]

using also \[\sum_{j=1}^{l} \cos(s \lambda_j) = O(n/s), \text{ see Zygmund (2002, p. 2). This bound is better when} \ s > n/m. \]

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 (33) is \(O(n^{-1})\). The second term of (33) 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 \(s \leq \lceil n/m^{2/3} \rceil\), the last sum is

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

and (33) = \(O(\ n^{-1} + m^{-1/3})\).

We still need to show that the mean of (31) is asymptotically equivalent to \(\sum_{i=1}^{p-1} \sum_{k=1}^{p-1} \eta_i \eta_k g_{pp} J_{ik}\).
Thus,

\[
E (31) = \sum_{t=1}^{n} \sum_{s=1}^{t-1} E \left( \text{tr} \left( c_{t-s,n}^t c_{t-s,n} e_s e'_s \right) \right) \\
= \sum_{t=1}^{n} \sum_{s=1}^{t-1} \text{tr} \left( c_{t-s,n}^t c_{t-s,n} \right) \\
= \sum_{t=1}^{n} \sum_{s=1}^{t-1} \sum_{j=1}^{m} \sum_{k=1}^{m} \frac{1}{4\pi^2 n^2 m} \text{tr} \left( \theta'_j \theta_k \right) \cos \left( (t-s) \lambda_j \right) \cos \left( (t-s) \lambda_k \right) \\
= \sum_{t=1}^{n} \sum_{s=1}^{t-1} \sum_{j=1}^{m} \sum_{k=1}^{m} \frac{1}{4\pi^2 n^2 m} \text{tr} \left( \theta'_j \theta_k \right) \cos^2 \left( (t-s) \lambda_j \right) + \sum_{t=1}^{n} \sum_{s=1}^{t-1} \sum_{j=1}^{m} \sum_{k=1}^{m} \frac{1}{4\pi^2 n^2 m} \text{tr} \left( \theta'_j \theta_k \right) \cos \left( (t-s) \lambda_j \right) \cos \left( (t-s) \lambda_k \right). 
\]  

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

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

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 can bound (35) by

\[
O \left( \sum_{j=1}^{m} \sum_{k \neq j}^{m} \left( n^2 m \right)^{-1} \frac{1}{n} \right) = O \left( m/n \right). \]  

Now, \( \text{tr} \left( \theta'_j \theta_j \right) \) is equal to

\[
\sum_{i,k=1}^{p-1} \text{tr} \left( \eta_i \eta_k \lambda_m^{d_i+d_k+2d_e} \text{Re} \left( A_p^* (\lambda_j) A_i (\lambda_j) + A_i^* (\lambda_j) A_p (\lambda_j) \right) \text{Re} \left( A_k^* (\lambda_j) \tilde{A}_p (\lambda_j) + A_p^* (\lambda_j) \tilde{A}_k (\lambda_j) \right) \right)
\]

\[
= \sum_{i,k=1}^{p-1} \eta_i \eta_k \lambda_m^{d_i+d_k+2d_e} 4\pi^2 \left( f_{pp} (\lambda_j) f_{ik} (\lambda_j) + f_{ip} (\lambda_j) f_{kp} (\lambda_j) + f_{pk} (\lambda_j) f_{pi} (\lambda_j) + f_{pp} (\lambda_j) f_{ki} (\lambda_j) \right)
\]

by definition of \( f (\lambda) \) in (11). The second and third terms are of smaller order than the first and fourth terms by Assumption A’, so (34) reduces to

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

28
since

(where so that we can rewrite (36) as

\( \frac{2\pi}{n} \sum_{j=1}^{m} (f_{pp}(\lambda_j) f_{ik}(\lambda_j) + f_{pp}(\lambda_j) f_{ki}(\lambda_j)) \sim \int_{0}^{\lambda_m} (f_{pp}(\lambda) f_{ik}(\lambda) + f_{pp}(\lambda) f_{ki}(\lambda)) d\lambda \)

\( \sim \int_{0}^{\lambda_m} (g_{ik} + g_{ki}) g_{pp} \lambda^{-d_i - d_k - 2d_e} d\lambda \)

\( = 2g_{ik}g_{pp} \frac{\lambda^{1-d_i-d_k-2d_e}}{1-d_i-d_k-2d_e}, \)

so that we can rewrite (36) as

\[
(36) \sim \sum_{i=1}^{p-1} \sum_{k=1}^{p-1} \eta_i \eta_k \left( \sum_{t=1}^{n} \sum_{s=1}^{t-1} \cos^2 ((t-s) \lambda_j) \right) \frac{1}{\pi m} g_{ik} g_{pp} \frac{\lambda_m}{1-d_i-d_k-2d_e},
\]

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

\[
(36) = \sum_{i=1}^{p-1} \sum_{k=1}^{p-1} \eta_i \eta_k \frac{1}{2} g_{ik} g_{pp} \frac{1}{1-d_i-d_k-2d_e} = \sum_{i=1}^{p-1} \sum_{k=1}^{p-1} \eta_i \eta_k g_{pp} J_{ik},
\]

where

\[
J_{ik} = \frac{g_{ik}}{2(1-d_i-d_k-2d_e)}, \quad 1 \leq i, k \leq p-1,
\]

and we have shown (28).

To show (30),

\[
\sum_{t=1}^{n} E \left( z_{tn}^4 \right) = \sum_{t=1}^{n} E \left( \sum_{s=1}^{t-1} c_{t-s,n} \varepsilon_{t} \varepsilon_{t}' \sum_{r=1}^{t-1} c_{t-r,n} \varepsilon_{r} \varepsilon_{t}' \sum_{p=1}^{t-1} c_{t-p,n} \varepsilon_{p} \varepsilon_{t}' \sum_{q=1}^{t-1} c_{t-q,n} \varepsilon_{q} \right)
\]

\[
\leq C \left( \sum_{t=1}^{n} \text{tr} \left( \sum_{s=1}^{t-1} c_{t-s,n} c_{t-s,n}' c_{t-s,n} c_{t-s,n}' \right) + \sum_{t=1}^{n} \text{tr} \left( \sum_{s=1}^{t-1} c_{t-s,n} c_{t-s,n}' \sum_{r=1}^{t-1} c_{t-r,n} c_{t-r,n}' \right) \right),
\]

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

Hence, we have shown that

\[
\sqrt{m} \lambda_m^{-1} \hat{F}_{xe} \overset{d}{\to} N(0, g_{pp} J)
\]

and the theorem follows using (19), (20), and (38) and applying the Slutsky Theorem. ■

29
References


Robinson, P. M. (1994a), ‘Rates of convergence and optimal spectral bandwith for long range


Robinson, P. M. (1995b), ‘Log-periodogram regression of time series with long range depen-


Robinson, P. M. & Hualde, J. (2003), ‘Cointegration in fractional systems with unkown inte-


Robinson, P. M. & Marinucci, D. (2003), Semiparametric frequency domain analysis of frac-

research’, *Advances in Futures and Options Research* **2**, 209–221.

tion’, *Cowles Foundation Discussion Paper 1367*.

Velasco, C. (2003), ‘Gaussian semi-parametric estimation of fractional cointegration’, *Journal
Table 1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Realized volatility</th>
<th>Implied volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>−2.642928</td>
<td>−2.035616</td>
</tr>
<tr>
<td>Variance</td>
<td>0.122568</td>
<td>0.200101</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.585426</td>
<td>−0.916736</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>0.457148</td>
<td>2.807389</td>
</tr>
</tbody>
</table>

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

Table 2: Fractional Integration Order

<table>
<thead>
<tr>
<th></th>
<th>Realized volatility</th>
<th>Implied volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$y_t = \ln \sigma_{RV,t}$</td>
<td>$x_t = \ln \sigma_{IV,t}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$y_t$</th>
<th>$x_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Box-Pierce (lags)$^a$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Q$ ($L = 4$)</td>
<td>394.7***</td>
<td>136.8***</td>
</tr>
<tr>
<td>$Q$ ($L = 24$)</td>
<td>1211**</td>
<td>649.4**</td>
</tr>
<tr>
<td>GSP (bandwidth)$^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{d}$ ($m = 20$)</td>
<td>0.48468</td>
<td>0.46281</td>
</tr>
<tr>
<td></td>
<td>(0.11180)</td>
<td>(0.11180)</td>
</tr>
<tr>
<td>$\hat{d}$ ($m = 37$)</td>
<td>0.44758</td>
<td>0.45273</td>
</tr>
<tr>
<td></td>
<td>(0.08220)</td>
<td>(0.08220)</td>
</tr>
<tr>
<td>$\hat{d}$ ($m = 68$)</td>
<td>0.41620</td>
<td>0.35033</td>
</tr>
<tr>
<td></td>
<td>(0.06063)</td>
<td>(0.06063)</td>
</tr>
<tr>
<td>ADF (lags)$^c$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t$ ($L = 0$)</td>
<td>−11.079**</td>
<td>−14.843**</td>
</tr>
<tr>
<td>$t$ ($L = 4$)</td>
<td>−4.6995**</td>
<td>−5.3166**</td>
</tr>
</tbody>
</table>

$^a$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 $\sqrt{m}(\hat{d} - d) \rightarrow_d N (0, 1/4)$.

$^c$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

<table>
<thead>
<tr>
<th>Bandwidth m</th>
<th>$\hat{\alpha}_m$</th>
<th>$\hat{\beta}_m$</th>
<th>s.e.$(\hat{\beta}_m)$</th>
<th>$\hat{d}_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = n - 1$</td>
<td>$-1.8628$</td>
<td>$0.38325$</td>
<td>$-0.38996$</td>
<td>$0.33443$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.11180)$</td>
<td>$(0.08270)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.31102)$</td>
<td>$(0.06063)$</td>
</tr>
<tr>
<td>$m = 15$</td>
<td>$-0.91075$</td>
<td>$0.85094$</td>
<td>$0.18866$</td>
<td>$0.17005$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$0.15844$</td>
<td>$0.10991$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.11180)$</td>
<td>$(0.08270)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.09479)$</td>
<td>$(0.06063)$</td>
</tr>
<tr>
<td>$m = 9$</td>
<td>$-0.93154$</td>
<td>$0.84072$</td>
<td>$0.22678$</td>
<td>$0.19860$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$0.18376$</td>
<td>$0.11400$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.11180)$</td>
<td>$(0.08270)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.09752)$</td>
<td>$(0.06063)$</td>
</tr>
<tr>
<td>$m = 6$</td>
<td>$-0.93049$</td>
<td>$0.84124$</td>
<td>$0.25163$</td>
<td>$0.21971$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$0.20286$</td>
<td>$0.11378$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.11180)$</td>
<td>$(0.08270)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.09738)$</td>
<td>$(0.06063)$</td>
</tr>
<tr>
<td>$m = 3$</td>
<td>$-0.83362$</td>
<td>$0.88882$</td>
<td>$0.25279$</td>
<td>$0.23727$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$0.21978$</td>
<td>$0.09856$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.11180)$</td>
<td>$(0.08270)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$(0.08669)$</td>
<td>$(0.06063)$</td>
</tr>
</tbody>
</table>

This table reports semiparametric estimates of the stationary fractional cointegration relation (18) 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 order $d_x$ of the raw regressor series, used in the calculation of standard errors of $\hat{\beta}_m$ based on (12), the results for bandwidth 68 from Table 2 were used.
Figure 1: Autocorrelation function up to lag 100 for log realized volatility
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 $m = 15$
List of D-CAF’s Working Papers
2006 -


2. B.J. Christensen and M. Ørregaard Nielsen (March 2006), The implied-realized volatility relation with jumps in underlying asset prices – pdf.file


6. B.J. Christensen and N.M. Kiefer (March 2006), Investment in advertising campaigns and search: Identification and inference in marketing and dynamic programming models – pdf-file