Unbalanced Fractional Cointegration and the No-Arbitrage Condition on Commodity Markets
Gilles De Truchis, Florent Dubois

To cite this version:
Gilles De Truchis, Florent Dubois. Unbalanced Fractional Cointegration and the No-Arbitrage Condition on Commodity Markets. 2014. <halshs-01065775>

HAL Id: halshs-01065775
https://halshs.archives-ouvertes.fr/halshs-01065775
Submitted on 18 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés. 
Unbalanced Fractional Cointegration and the No-Arbitrage Condition on Commodity Markets

Gilles de Truchis
Florent Dubois
Unbalanced Fractional Cointegration and the No-Arbitrage Condition on Commodity Markets

Gilles de Truchis¹,∗, Florent Dubois¹

Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS

Preliminary draft
September 2014

Technical and non-technical abstracts

A necessary condition for two time series to be nontrivially cointegrated is the equality of their respective integration orders. Nonetheless, in some cases, the apparent unbalance of integration orders of the observables can be misleading and the cointegration theory applies all the same. This situation refers to unbalanced cointegration in the sense that balanced long run relationship can be recovered by an appropriate filtering of one of the time series. In this paper, we suggest a local Whittle estimator of bivariate unbalanced fractional cointegration systems. Focusing on a degenerating band around the origin, it estimates jointly the unbalance parameter, the long run coefficient and the integration orders of the regressor and the cointegrating errors. Its consistency is demonstrated for the stationary regions of the parameter space and a finite sample analysis is conducted by means of Monte Carlo experiments. An application to the no-arbitrage condition between crude oil spot and futures prices is proposed to illustrate the empirical relevance of the developed estimator.

The no-arbitrage condition between spot and future prices implies an analogous condition on their underlying volatilities. Interestingly, the long memory behavior of the volatility series also involves a long-run relationship that allows to test for the no-arbitrage condition by means of cointegration techniques. Unfortunately, the persistent nature of the volatility can vary with the future maturity, thereby leading to unbalanced integration orders between spot and future volatility series. Nonetheless, if a balanced long-run relationship can be recovered by an appropriate filtering of one of the time series, the cointegration theory applies all the same and unbalanced cointegration operates between the raw series. In this paper, we introduce a new estimator of unbalanced fractional cointegration systems that allows to test for the no-arbitrage condition between the crude oil spot and futures volatilities.

Keywords: Unbalanced cointegration, Fractional cointegration, No-arbitrage condition, Local Whittle likelihood, Commodity markets

JEL: C22, G10

1. Introduction

This paper deals with the estimation of a general class of models, terms unbalanced cointegration systems, that encompass the well known triangular cointegration system. The former systems originate from very recent developments (see Hualde 2006), while the latter relies on the seminal definition of

∗Corresponding author

Email addresses: gilles.de-truchis@univ-amu.fr (Gilles de Truchs), florent.dubois@univ-amu.fr (Florent Dubois)

¹Aix-Marseille School of Economics, Aix-Marseille University, Château Lafarge, Route des Milles 13290 Aix-en-Provence, France. Tel.: 00 33 (0)4 42 93 59 60; fax: 00 33 (0)4 42 38 95 85.
the cointegration (see Granger 1981) which establishes that two time series $y_t$ and $x_t$ share a common stochastic trend, if (i) $y_t$ and $x_t$ are both integrated of order $\delta_2$, hereafter $I(\delta_2)$; (ii) there exists a non-null scalar $\beta$ so that $e_t = y_t - \beta x_t \sim I(\delta_1)$ and $\delta_2 - \delta_1 > 0$. This definition is quite general and does not constrain integration orders to be integers. Nonetheless, following the paper of Engle and Granger (1987), the literature has primarily investigated the particular case where observables are unit root processes, i.e. $I(1)$, and a linear combination between them has short memory, i.e. $I(0)$.

In a pioneer work, Cheung and Lai (1993) extended the model of Engle and Granger (1987), allowing the integration order of the cointegration errors (i.e. $\delta_1$) to be a real number. Their methodology is simple and operates in two steps if $y_t$ and $x_t$ are $I(1)$. When $y_t$ and $x_t$ are fractionally integrated (i.e. $\delta_2 \in \mathbb{R}$), some new complications arise. While a part of the literature has focused on testing the homogeneity of integration orders between $y_t$ and $x_t$ (see e.g Robinson and Yajima 2002, Nielsen and Shimotsu 2007, Hualde 2013), another part has investigated the estimation of such systems with unknown integration orders (see e.g Robinson and Hualde 2003). In the latter case, many difficulties appear in the uniform treatment of the objective functions on the entire parameter space. Thereby, three cases are generally distinguished in the literature. Considering that a fractionally integrated process is stationary when the integration order is less than 1/2 and nonstationary otherwise, and interpreting $\delta_2 - \delta_1$ as the cointegration strength, (i) the strong cointegration occurs when $\delta_2 - \delta_1 > 1/2$; (ii) the weak cointegration occurs when $\delta_2 - \delta_1 < 1/2$; (iii) the stationary cointegration occurs when $\delta_1 < \delta_2 < 1/2$. As Velasco (2003) suggested, it is more efficient to estimate simultaneously all the parameters of interest (see also Lobato 1999). Among recent contributions in this direction we can mention Hualde and Robinson (2007) and Shimotsu (2012) concerning the weak and strong cointegration cases and Nielsen (2007) and Robinson (2008) for the stationary case. In this paper we are particularly interested in the stationary cointegration case which is mainly attractive in empirical finance where time series have long range dependence but are likely to be stationary (e.g. volatility, volume, closing prices of commodities). Investigating stationary cointegration is also of interest because spurious regression can occur, whether $y_t$ and $x_t$ are stationary or not, as long as their integration orders sum up to a value greater than 1/2 (see Tsay and Chung 2000). Nonetheless, it is not so easy to identify whether or not the observed time series are stationary leading practitioners to possibly make a wrong decision. Accordingly, we are also interested in non-stationary region of the parameter space, although the proposed estimator is not theoretically designed to handle this case. In the present paper, this issue is investigating by means of Monte Carlo study.

Meanwhile, Hualde (2006) investigated a promising alternative avenue of research termed unbalanced cointegration. Let $y_t$ and $x_t$ be two observable time series integrated of orders $\delta_2$ and $\delta_2 + \xi$ respectively. Following Hualde (2006), unbalanced cointegration is likely to occur between $y_t$ and $x_t$ and cointegration theory, in the usual sense, is likely to apply between $y_t$ and $x_t(\xi)$, if there exists a linear combination.

---

2See Granger (2010) for a very simple and short introduction to the cointegration theory.

3See also Dueker and Startz (1998) and de Truchis (2013) for a fully parametric approach in time and frequency domain respectively. Hualde and Robinson (2010) and Johansen and Nielsen (2012) deal with extensions to multivariate case in frequency and time domain respectively, but this discussion goes beyond of the scope of the paper.
between them which has less memory $\delta_1$. Accordingly, this type of cointegration does not differ from the original cointegration theory but is useful from an empirical point of view. Hualde (2006) suggests to estimate $\xi$ by the difference between integration orders of $y_t$ and $x_t$ and discusses the consistency of the OLS estimator of $\beta$. With respect to the joint estimation of $\beta$ and $\xi$, the standard OLS are unfeasible and Hualde (2014) investigated the consistency of the non-linear least squares estimates of $\xi$ and $\beta$ when $\delta_2 > 1/2$ and $\delta_1 \geq 0$. Conversely, we focus here on the stationary case assuming that $\delta_2 \in (0, 1/2 - |\xi|)$ and $\delta_2 > \delta_1 \geq 0$. We propose a local Whittle estimator of bivariate unbalanced fractional cointegration systems. Focusing on a degenerating band around the origin of the spectral density matrix, it estimates jointly the unbalance parameter, the long run coefficient and the integration orders of the regressor and the cointegrating errors without specifying the short run dynamics avoiding thereby the misspecification issue. The consistency of the estimator is discussed and highlight that $\beta$ is $(m/n)^{\delta_1 - \delta_2}$-consistent when $x_t$ is appropriately filtered in the long run equation, with $m$ the bandwidth number and $n$ the sample size. The finite sample properties are also investigated by means of Monte Carlo simulations for a wide range of specifications.

In the empirical part of the paper, we aim to test whether the no-arbitrage condition holds on the crude oil market. A naive approach would be to investigate the presence of cointegration between spot and future prices. But as argued by Brenner and Kroner (1995), they should not be cointegrated because commodity markets are subject to the so-called convenience yield, that is probably not a short memory process. Consequently the long-run equation is contaminated by an additive persistent component. However, as demonstrated by Liu and Tang (2010), the no-arbitrage condition can exist on commodity market if the convenience yield is non-negative and the no-arbitrage issue remains unsolved. Recently, Rossi and Santucci de Magistris (2013), have strengthened that the no-arbitrage condition between spot and future asset prices implies an analogous condition on their underlying volatilities. This original approach is convenient because it can be adapted to commodity markets under mild conditions on the volatility of the convenience yield. Accordingly, we aim to test for the no-arbitrage condition between the crude oil spot and futures volatilities rather than spot and future prices. The unbalanced cointegration framework is particularly appropriated here because we find some evidence that the persistent nature of the volatility can vary with the future maturity (i.e. unbalanced integration orders between spot and future volatility series).

The rest of the paper is laid out as follows. In Section 2, we introduce a bivariate stationary model of unbalanced cointegration. In Section 3, we develop the local Whittle estimator of unbalanced fractional cointegration (LWE-UFC). In Section 4 we demonstrate the consistency of the proposed estimator. Finite sample properties are investigated in Section 5. The application to the no-arbitrage condition between the volatilities of the spot and futures prices is proposed in Section 6. The Section 7 concludes the paper. Proofs are given in 8. Additional results and simulations are reported in 10 and 11.

---

4 Hualde (2014) interestingly finds that the limiting distributions of estimates depend on a modified version of the Type II fractional Brownian motion. Different properties of this new type of fractional Brownian motion are discussed in Hualde (2012).
2. A stationary model of unbalanced cointegration

In the following, we say that a stochastic process \( \xi_t \), has long memory \( \alpha \in (0,1/2) \), if its spectral density \( f_\xi(\lambda) \) satisfies \( f_\xi(\lambda) \sim g\lambda^{-2\alpha} \) as the frequency \( \lambda \) tends to 0, where the notation \( \sim \) means that the ratio of the left and right sides tends to 1 in the limit. Then, \( \xi_t \) has short memory when \( \alpha = 0 \) and intermediate memory when \( \alpha \in (-1/2, 0) \).\(^5\) Now, we consider an unbalanced bivariate form of the triangular system introduced in Phillips (1991) and extended to the fractional framework by Nielsen (2004). Let \( y_t \) and \( x_t \) be two unbalanced observable variables with unknown real integration orders, \( \delta_2 \) and \( \delta_2 + \xi_n \) respectively. Hualde (2006) states that \( y_t \) and \( x_t \) are weakly unbalanced when \( \delta_2 \) and \( \delta_2 + \xi_n \) does not diverge at infinity (i.e. \( \xi_n \to 0 \) as \( n \to \infty \)) and strongly unbalanced when \( \xi_n = |\xi| > 0 \) as \( n \to \infty \). To simplify notation, we shall use \( \zeta \) to denote \( \xi_n \). Then, the triangular unbalanced cointegration system is defined by

\[
y_t = \beta x_t(\xi) + u_{1t}(-\delta_1), \quad x_t = u_{2t}(-\delta_2 - \xi), \quad t = 1, 2, ..., n, \tag{1}
\]

where generically, \((-a)\) denotes the fractional filter \( (1 - L)^{-a} = \sum_{k=0}^{\infty} a_k(\alpha) L^k \) with \( a_k(\alpha) := \Gamma(k + a)/(\Gamma(a)k!)^{-1} \), \( L \) the lag operator and \( \Gamma(\cdot) \), the gamma function. Although standard cointegration theory does not apply to \( y_t \) and \( x_t \) , it does to \( y_t \) and \( x_t(\xi) \). Thereby, the System (1) is a cointegration system in the sense that both series have a dominant common component with memory \( \delta_2 \) that can be suitably recovered by filtering \( x_t \sim I(\delta_2 + \zeta) \) to obtain \( x_t(\xi) \sim I(\zeta) \).

**Assumption 1.** \( y_t, x_t \) and \( y_t - \beta x_t(\xi) \) are covariance stationary processes integrated of orders \( \delta_2, \delta_2 + \zeta \) and \( \delta_1 \) respectively, and satisfying

\[
0 \leq \delta_1 < \delta_2 < |\zeta| < 1/2 \tag{2}
\]

where \( |\zeta| < k \), with \( k \) an arbitrary real number small compared to \( \delta_2 \).

Under Assumption 1, the System in (1) provides a valid data-generating process for stationary cointegration model so that \( z_t = (y_t - \beta x_t(\xi), x_t)' \) possesses a spectral density, \( f_z(\lambda_j) \), where \( \lambda_j \) denotes the Fourier frequencies, \( \lambda_j = 2\pi j/n \), with \( j = 1, \ldots, m \) and \( m = o(n) \), the bandwidth parameter. Assumption 1 leaves out anti-persistent processes because they clearly have limited economic relevance.

Now, assume that \( u_t = (u_{1t}, u_{2t})' \) has short memory with spectral density \( f_u(\lambda_j) \) satisfying, \( f_u(\lambda_j) \sim G \) in the neighborhood of the origin, with \( G \) (the long run covariance matrix) a real, symmetric, finite and positive definite matrix. Notice that when cointegration arises, \( \text{rank}(G) < 2 \), so that \( G \) has reduced-rank whether or not \( \zeta \neq 0 \) (see Hualde 2006).

\(^5\)When \( \alpha \in [1/2,1] \), the spectral density of \( \xi_t \) is no longer defined although Velasco (1999b) demonstrated it has a pseudo-spectral density and standard local Whittle-based estimators are biased. In such a case, Velasco (1999b) proposed to use a tapered periodogram to reduce the bias, but at the cost of a higher variance. Exploiting this result, Shimotsu (2012) combined a multivariate extension of the exact local Whittle estimator of Shimotsu (2010) and a tapered version of the Robinson (2008) to propose a fractional cointegration estimator which is consistent over \(-1/2 < \delta_1 < \delta_2 < \infty \) and includes \( \delta_2 = 1 \) and \( \delta_1 = 0 \) as a special case.
Remark 1. Assumptions on \( u_t \) entails only mild conditions so that \( u_t \) is possibly a vector ARMA process or any other short memory processes with a Wold representation, \( u_t = C(L)e_t \), where \( e_t \) are further defined as martingale difference innovations and \( C(L) \) is a square-summable causal matrix filter satisfying \( G = C(1)C(1)'/(2\pi)^{-1} \).

In the present study, we define processes only on the vicinity of the origin, in view of allowing a semi-parametric treatment of the short-run dynamics. Indeed, we support that such approach is particularly of interest in cointegration analysis because empirical interest is more likely to lie in long run dynamics. Accordingly, by Assumption 1 and considering the Remark 1, \( z_t \) has a spectral density, so that

\[
E(z_t - E(z_t)) (z_{t+k} - E(z_t')) = \int_{-\pi}^{\pi} e^{ik\lambda} f_{z}(\lambda) d\lambda,
\]

with \( f_{z}(\lambda) = \Lambda(\lambda)^{-1}f_{u}(\lambda) (\Lambda(\lambda)^{*})^{-1} \) and \( \Lambda(\lambda) = \text{diag} \left( (1 - e^{i\lambda})^{\delta_1}, (1 - e^{i\lambda})^{\delta_2 + \xi} \right) \). Since \( (1 - e^{i\lambda})^{a} = (|2 \sin(\lambda/2)|)^{a} \sim \lambda^{a} \) as \( \lambda \to 0 \), we can avoid a parametric treatment of \( f_{z}(\lambda) \) in favor of the following local power law representation around zero frequency,

\[
f_{z}(\lambda) \sim \left( \Lambda(\lambda; \theta_1) \right)^{-1} G \left( \Lambda(\lambda; \theta_1)^{*} \right)^{-1}, \quad \Lambda(\lambda; \theta_1) = \text{diag} \left( \lambda^{\delta_1}, \lambda^{\delta_2 + \xi} \right), \quad \text{as } \lambda \to 0
\]

where \( \theta_1 = (\delta_1, \delta_2 + \xi)' \) and the superscript * denotes the conjugate transpose. As Nielsen (2007) we assume that \( G \) is diagonal so that \( u_{11} \) and \( u_{21} \) are incoherent in the vicinity of the origin and the Equation (4) is correctly specified.\(^6\) Thereby, the phase parameter modeled as \( \varphi = (\pi - \lambda)(\delta_2 - \delta_1)/2 \) in Robinson (2008) and Shimotsu (2012) is null in our framework (see also Shimotsu 2007).\(^7\) Notice also that, if \( \delta_2 \leq \delta_1 \), \( \beta \) cannot be identified. This issue also occurs in standard regression analysis of balanced cointegration.

A matrix representation of the System (1) gives

\[
\begin{pmatrix}
1 & -\beta(1-L)^{\xi} \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
y_t \\
x_t
\end{pmatrix}
= \begin{pmatrix}
(1-L)^{-\delta_1} & 0 \\
0 & (1-L)^{-\delta_2 - \xi}
\end{pmatrix}
\begin{pmatrix}
u_{1t} \\
u_{2t}
\end{pmatrix}.
\]

The local Whittle analysis of restricted versions of Equation (5) is not new. For instance, when \( \beta = 0 \), \( y_t \) and \( x_t \) do not share any long-run component, and the first matrix on the left side of (5) reduces to the identity matrix. Thereby, one can recover the so-called stationary ARFIMA model whose the local Whittle estimate has been studied in univariate framework by Robinson (1995a) and later extended to a multivariate setting by Lobato (1999) and Shimotsu (2007). When \( \beta \neq 0 \) and \( \xi = 0 \), cointegration can arise in the usual sense. Considering this latter case, Robinson and Marinucci (2003) and Nielsen (2005) discuss the estimation of \( \beta \) by means of frequency domain least squares (see also Robinson 1994, Christensen and Nielsen 2006, Nielsen and Frederiksen 2011). As mention previously, approximate local


\(^7\)The presence of non-null off-diagonal elements in \( G \) should imply non-negligible imaginary part of the cross-spectrum element \( f^{ab}_{z}(\lambda) \) such as \( f^{ab}_{z}(\lambda) \sim G_{ab}\lambda^{-\delta_1-\delta_2}e^{i(\pi - \lambda)(\delta_1 - \delta_2)/2} \) as \( \lambda \to 0 \), for \( a,b = 1,2 \) and where \( G_{ab} \) denotes the \((a,b)\)th element of \( G \).
Whittle estimation of $\delta_1, \delta_2$ and $\beta$ is considered in Nielsen (2007), Robinson (2008) and Shimotsu (2012). When $\xi \neq 0$ and $\delta_2 > 1/2$, Hualde (2014) derives the asymptotic properties of the non-linear least squares estimator of $\beta$ and $\xi$.

3. Local Whittle estimation

In the following, we introduce a local Whittle estimator of $\theta = (\delta_1, \delta_2, \xi, \beta)'$. Let, $I_z$ be the periodogram matrix of $z_t$, defined as $I_z(\lambda_j; \theta_2) = w_z(\lambda_j; \theta_2) w_z(\lambda_j; \theta_2)^*$ with $w_z(\lambda_j; \theta_2) = (2\pi n)^{-1/2} \sum_{t=1}^{n} z_t e^{it\lambda_j}$, the Fourier transform of $z_t$ and $\theta_2 = (\beta, \xi)'$. Updated to bandwidth $m = o(n)$, i.e. $j = 1, \ldots, m$, in view of the local treatment we obtain,

$$ I_z(\lambda_j; \theta_2) = \begin{pmatrix} w_y(\lambda_j) - \beta \hat{\lambda}_j^2 w_x(\lambda_j) \\ w_x(\lambda_j) \end{pmatrix} \begin{pmatrix} w_y(\lambda_j) - \beta \hat{\lambda}_j^2 w_x(\lambda_j) \\ w_x(\lambda_j) \end{pmatrix}^* . \quad (6) $$

Thereby, the presence of $\hat{\lambda}_j^2$ corrects for the fact that long memory parameters of $y_t$ and $x_t$ are unbalanced. Then, the discrete local Whittle approximation to the likelihood is

$$ Q_m(\theta, G) = m^{-1} \sum_{j=1}^{m} \left[ \log \det \left( \left( \Lambda(\lambda_j; \theta_1) \right)^{-1} G \left( \Lambda(\lambda_j; \theta_1)^* \right)^{-1} \right) + \text{tr} \left( G^{-1} \Lambda(\lambda_j; \theta_1) I_z(\lambda_j; \theta_2) \Lambda(\lambda_j; \theta_1)^* \right) \right] , \quad (7) $$

where $G \in \Theta_G$, the set of real positive definite $2 \times 2$ matrices. The objective function $Q$ is minimized over $\Theta_G$ by

$$ \hat{\theta} = \text{arg min}_{\theta \in \Theta} Q_m(\theta, G) . \quad (8) $$

leading to the following concentrated likelihood function

$$ R_m(\theta) = \log \det \hat{\theta} - \frac{2(\delta_1 + \delta_2 + \xi)}{m} \sum_{j=1}^{m} \log \lambda_j . \quad (9) $$

Accordingly, the local Whittle estimator of $\theta$ is defined as $\hat{\theta} = \text{arg min}_{\theta \in \Theta} R_m(\theta)$, for $m \in [1, n/2]$ and $\Theta$ a compact subset of $\mathbb{R}^4$ with $\Theta = \Theta_{\delta} \times \Theta_{\xi} \times \Theta_{\beta}$ and $\delta = (\delta_1, \delta_2)'$. Vectors $\theta$ and $\hat{\theta}$ are respectively the vector of unknown and estimated values, $(\delta_1, \delta_2, \xi, \beta)'$ and $(\hat{\delta}_1, \hat{\delta}_2, \hat{\xi}, \hat{\beta})'$. Observe that Equation (8) yields $\hat{\theta}_{11}(\theta) = \text{Re} \left( m^{-1} \sum_{j=1}^{m} \lambda_j^2 I_z(\lambda_j; \theta_2) \right)$ with $I_z(\lambda_j; \theta_2) = I_{yy}(\lambda_j) - 2\beta \hat{\lambda}_j^2 I_{xy}(\lambda_j) + \beta^2 \hat{\lambda}_j^2 I_{xx}(\lambda_j)$, which has some similarities with the weighted least squares of Nielsen (2005). In 10, we show that the Proposition 1 of Nielsen (2005, p. 297) and the Theorem 1 of Robinson and Marinucci (2003) remain valid when $\xi \neq 0$ if $x_t$ is appropriately differenced. Accordingly, we anticipate that $\hat{\beta}$ is also $\lambda_m^{\beta} - \delta_2$-consistent when $\xi \neq 0$. 


4. Consistency

To prove the consistency of this local Whittle estimator, we introduce several assumptions, fairly similar to those of Shimotsu (2007) and Nielsen (2007). In the following, \( \theta_0 \) and \( G_0 \) will denote the true parameter values of \( \theta \) and \( G \). Then, let \( f_{abz}(\lambda) \) and \( G_{ab}^0 \) denote the \((a,b)\)th element of \( f_z(\lambda) \) and \( G_0 \) respectively. Define also \( \vartheta_01 = (\delta_{01}, \delta_{02} + \xi_0)' \) and \( \delta_a \) the \( a \)th element of \( \vartheta_01 \).

**Assumption 2.** As \( \lambda \to 0^+ \), elements of the spectral density \( f_z(\lambda) \) satisfies

\[
f_{abz}(\lambda) = G_{ab}^0 \lambda^{-\delta_{ab}} + o(\lambda^{-\delta_{ab}}), \quad a, b = \{1, 2\}, \tag{10}
\]

where matrix \( G \) is finite, real, symmetric and positive definite. Also assume \( G_{12} = G_{21} = 0 \).

**Assumption 3.** The sequence \( z_t \) is a linear process defined as

\[
z_t - E(z_t) = A(L)e_t = \sum_{j=0}^{\infty} A_j e_{t-j}, \quad \sum_{j=0}^{\infty} ||A_j||^2 < \infty, \tag{11}
\]

with \( ||.|| \) the Euclidean norm, so that \( A_j \) is a causal square summable matrix filter. Moreover, \( e_t \) satisfies, almost surely, \( E(e_t|F_{t-1}) = 0 \), \( E(e_t e'_t|F_{t-1}) = I_2 \), with \( F_t \) a \( \sigma \)-field generated by \( \{e_s, s \leq t\} \) and there exist a random variable \( \epsilon \) such that \( E(\epsilon^2 < \infty) \) and for all \( \eta > 0 \) and some \( K > 0 \), \( \Pr(||\epsilon||^2 > \eta) \leq K \Pr(\epsilon^2 > \eta) \).

Assumption 2 implies a zero-coherence condition that applies only in the vicinity of the origin. As argued in Nielsen (2007), it is a less restrictive assumption than the traditional orthogonality condition encountered in the least squares theory. Notably, it allows for errors to be correlated away from the origin and share, for instance, a common short- and/or medium-term dynamics. As mentioned previously, Robinson (2008) and Shimotsu (2012) relax this hypothesis. The present estimator could be modified to model \( f_z \) correctly when \( G^{12} \) is non-null by specifying the phase parameter and presumably, its consistency could be demonstrated in this case but we do not pursue that possibility further. Assumption 3 imposes uniformly square integrable martingale-difference innovations with constant conditional variance in view of the application of the standard CLT for martingale-difference arrays.\(^8\) The latter assumption on conditional variance could also probably be relaxed assuming boundedness of higher moments as in Robinson and Henry (1999) but we do not investigate this issue further.

**Assumption 4.** In a neighborhood of the origin, \( A(\lambda) = \sum_{j=0}^{\infty} A_j e^{ij\lambda} \) is differentiable and

\[
\frac{\partial}{\partial \lambda} A_a(\lambda) = O(\lambda^{-1}||A_a(\lambda)||) \text{ as } \lambda \to 0^+ \tag{12}
\]

where \( A_a(\lambda) \) is the \( a \)th row of \( A(\lambda) \).

\(^8\)We aim to investigate the asymptotic distribution of the estimator but for now the paper only discusses the consistency.
Assumption 4 implies \( \partial A_a(\lambda) / \partial \lambda = O(\lambda^{-\delta_1-1}) \) because by the Cauchy inequality

\[
||A_a(\lambda)|| \leq (A_a(\lambda)A_a^*(\lambda))^{1/2} = (2\pi f_{aa}(\lambda))^{1/2}.
\]

Thereby, under Assumption 3 and 4 we have \( f_z(\lambda) = (2\pi)^{-1}A(\lambda)A^*(\lambda) \).

**Assumption 5.** As \( n \to \infty \), the bandwidth parameter satisfies

\[
\frac{1}{m} + \frac{m}{n} \to 0.
\]

where \( m = \lfloor nk \rfloor \), \( k \in (0, 4/5] \).

The bandwidth requirement defined by the Assumption 5 ensures that \( m \) tends to \( \infty \) as \( n \to \infty \) but at a slower rate to remain in a neighborhood of the origin. The bandwidth parameter \( m \) is theoretically bounded by \( n^{4/5} \) but in practice a too small bandwidth increases the variance of the estimator while a too large \( m \) generally increases the bias. Accordingly, Assumptions 2-5 are analogous to Assumptions 1-4 of Shimotsu (2007), Nielsen (2007) and natural multivariate extensions of Assumptions A1-A4 of Robinson (1995a). Under Assumptions 2-5 we may now state the following theorem which establishes the convergence rate of \( \hat{\theta} \).

**Theorem 1.** Let Assumption 1-5 hold. Define \( \nu_0 = \delta_{02} - \delta_{01} \). Then, for \( \theta_0 \in \Theta \) and \( \delta_{01} < \delta_{02} \leq \delta_{02} + |\xi_0| \), as \( n \to \infty \),

\[
\begin{align*}
\left( \hat{\delta}_1 - \delta_{01} \right) &= O_p(m^{-1/2}) \\
\left( \hat{\delta}_2 - \delta_{02} \right) &= O_p(m^{-1/2}) \\
(\hat{\xi} - \xi_0) &= O_p(m^{-1/2}) \\
(\hat{\beta} - \beta_0) &= O_p(m^{-1/2}(n/m)^{-\nu_0})
\end{align*}
\]

The convergence rate of \( \hat{\beta} \) confirms that the system is rebalanced when \( x_t \) is appropriately filtered. Notice that when \( \delta_1 \to 0 \) and \( \delta_2 \to 1/2 \), \( \hat{\beta} \) is almost \( \sqrt{n} \)-consistent because \( m^{1/2}m^{\delta_1-\delta_2-1/2}n^{\delta_2-\delta_1-1/2}(\hat{\beta} - \beta_0) \overset{p}{\longrightarrow} 0 \). More importantly, when the presence of \( \xi \neq 0 \) is neglected and thus \( x_t(\xi) \) is trivially replaced by \( x_t \) in Equation (1), \( \hat{\beta} - \beta_0 = O_p((n/m)^{\delta_1-\delta_2-\delta}) \) and \( \hat{\beta} \) is likely to be inconsistent if \( \xi < 0 \).

5. Monte Carlo experiment

5.1. Simulation design

This section discusses the finite sample performance of the proposed estimator by means of Monte Carlo simulations. As argued in Hurvich and Ray (1995), it is not so easy for practitioners to identify
whether or not the observed time series are stationary. We deal with this issue by considering a data generating process (DGP) that accommodates the mean-reverting non-stationary regions of the parameter space, although the developed estimator is not theoretically designed to handle this case. Thereby, we generate a fractionally cointegrated system according to the following model,

\[ y_t = \beta x_t(\xi) + u_{1t}^\#(-\delta_1), \quad x_t = u_{2t}^\#(-\delta_2 - \xi), \quad t = 1, 2, ..., n, \]  

where for a generic process \( \zeta_t, \zeta_t^\# = \zeta_t l(t \geq 1) \), with \( l(.) \) the indicator function. \( \zeta_t^\#(-\alpha) \) denotes the fractional truncated filter \( \zeta_t^\#(-\alpha) := \sum_{k=0}^{-1} \Gamma(k + \alpha)(\Gamma(\alpha)k!)^{-1}\zeta_{t-k} \). The System in (17) provides a valid DGP for both, stationary and non-stationary regions of the parameter space, given that \( y_t \) and \( x_t \) are type II processes, thereby contrasting with our system in (1) which is of type I. Marinucci and Robinson (1999) shown that type I and type II processes are asymptotically equivalent in the stationary regions of the parameter space if both are generated from the same short memory sequence (see also Robinson 2005). Accordingly, the type II representation is often retained in simulation study for its simplicity. Nonetheless, Davidson and Hashimzade (2009) demonstrated that both representations can substantially differ in finite sample. We account for their recommendations by performing some additional simulations.
with a type I representation. Davidson and Hashimzade (2009) discuss the benefits and limitations of several techniques devoted to type I simulation. Given that we are interested in generating simple linear fractional processes, we use the circulant embedding method extended to multivariate fractional Gaussian noise by Helgason et al. (2011).

In our experiment we consider the following settings. The stationary cointegration case is explored for $\delta_2 = 0.4$ and $\delta_1 \in \{0.0, 0.2, 0.3\}$. Similarly, the strong and weak cointegration cases, are investigated for $\delta_1 = \{0.0, 0.2, 0.4\}$ and $\delta_1 = \{0.4, 0.6, 0.8\}$ respectively, with $\delta_2 = \{0.6, 0.8, 1.0\}$. Because in practice the weakly unbalanced cointegration case is generally indistinguishable from that of balanced cointegration, we do not consider this case in the simulation. Conversely, the strongly unbalanced cointegration case has greater applicability and is investigated for $\xi = \{-0.1, 0.1\}$. The long-run coefficient is fixed as $\beta = 1$. The vector $u_t$ is generated from a bivariate normal distribution $N_2(\mu, \Sigma)$ distribution. For each simulation, we report the bias, the variance and the Root Mean Squared Error (RMSE), defined by $\sqrt{I \sum_{i=1}^I (\hat{\theta}_i - \theta)^2}$ with $I$ the number of replications set to 10000. Following Shimotsu (2012), we use a penalty term defined as $\Pi(\beta, \hat{\beta}) = \min (0, \beta - \hat{\beta} + C) + \max (0, \beta - \hat{\beta} - C)$ to govern the objective function

Table 2: Simulation results for the stationary model when $\xi = 0.1$ and $\rho = 0.4$

<table>
<thead>
<tr>
<th>$m = [n^{0.5}]$</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_2$</td>
<td>$\delta_1$</td>
<td>$\hat{\beta}$</td>
<td>Bias</td>
</tr>
<tr>
<td>0.4</td>
<td>0</td>
<td>$\delta_2$</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_1$</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\xi$</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>0.345</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>$\delta_2$</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_1$</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\xi$</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>0.366</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3</td>
<td>$\delta_2$</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta_1$</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\xi$</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>0.357</td>
</tr>
</tbody>
</table>

with a type I representation. Davidson and Hashimzade (2009) discuss the benefits and limitations of several techniques devoted to type I simulation. Given that we are interested in generating simple linear fractional processes, we use the circulant embedding method extended to multivariate fractional Gaussian noise by Helgason et al. (2011).

In our experiment we consider the following settings. The stationary cointegration case is explored for $\delta_2 = 0.4$ and $\delta_1 \in \{0.0, 0.2, 0.3\}$. Similarly, the strong and weak cointegration cases, are investigated for $\delta_1 = \{0.0, 0.2, 0.4\}$ and $\delta_1 = \{0.4, 0.6, 0.8\}$ respectively, with $\delta_2 = \{0.6, 0.8, 1.0\}$. Because in practice the weakly unbalanced cointegration case is generally indistinguishable from that of balanced cointegration, we do not consider this case in the simulation. Conversely, the strongly unbalanced cointegration case has greater applicability and is investigated for $\xi = \{-0.1, 0.1\}$. The long-run coefficient is fixed as $\beta = 1$. The vector $u_t$ is generated from a bivariate normal distribution $N_2(\mu, \Sigma)$ distribution. For each simulation, we report the bias, the variance and the Root Mean Squared Error (RMSE), defined by $\sqrt{I \sum_{i=1}^I (\hat{\theta}_i - \theta)^2}$ with $I$ the number of replications set to 10000. Following Shimotsu (2012), we use a penalty term defined as $\Pi(\beta, \hat{\beta}) = \min (0, \beta - \hat{\beta} + C) + \max (0, \beta - \hat{\beta} - C)$ to govern the objective function

To generate nonstationary series (e.g. $x_t$ with $\delta_2 \geq 1/2$), we simulate an intermediate stationary (or asymptotically stationary) process, integrated of order $I(\delta_2 = \delta_2 - 1)$ and cumulate the resulting series.
when the space for dimensionality reduction is weak, with \( \hat{\beta} = \hat{\beta}_{LSE} \), so that it preserves the asymptotic results obtained in Equation (2). Alternatively, the Narrow-Band Least Squares (NBLS) estimate has also been used for initialization and does not significantly modify the results. The initial estimates for \( \hat{\delta}_x \) and \( \hat{\delta}_y \) are obtained from the local Whittle estimator of Robinson (1995a) applied to \( x_t \) and \( y_t \) respectively. Therefore, the initial estimate for \( \hat{\xi} \), namely \( \hat{\xi}_{LSE} \), results from the regression of \( y_t \) and \( x_t(\hat{\xi}) \). All computations are performed using MATLAB 2013a.

5.2. Simulation results

Table 1 reports the results for 10000 replications of the stationary cointegration model when \( \zeta = 0.1 \) and \( \rho = 0 \). Accordingly, this specification satisfies all the assumptions of the model defined in Equation (1). In absence of short-run dynamics, taking frequencies away from the origin essentially impact the variance (rather than the bias) compare with \( m = \lfloor n^{0.5} \rfloor \). Thereby, in both cases, the approximation in Equation (4) is close to \( f_z(\lambda) = (2\pi)^{-1} A(\lambda) A(\lambda)^* \). Observe that the strength of the cointegration and the variance of \( \hat{\beta} \) are negatively linked which is in accordance with the limit theory (the convergence rate of \( \hat{\beta} \) depends on the strength of the cointegration). In all cases, the variances (the RMSE so on) decrease as the sample size increases. Moreover, variances are lower when \( m = \lfloor n^{0.8} \rfloor \).

Table 2 reports the results for 10000 replications of the stationary cointegration model when \( \zeta = 0.1 \) and \( \rho = 0.4 \). Thereby, the off-diagonal elements of \( G \) are non-null and a correlation between \( y_t - \beta x_t(\hat{\xi}) \)
Table 3: Comparison of type I and II processes when $\xi = -0.1$, $\rho = 0$ and $m = \lceil n^{0.8} \rceil$.

<table>
<thead>
<tr>
<th>Type I</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>SD</td>
<td>RMSE</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>-0.059</td>
<td>0.029</td>
<td>0.180</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.045</td>
<td>0.017</td>
<td>0.138</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.136</td>
<td>0.015</td>
<td>0.183</td>
</tr>
<tr>
<td>$\hat{\theta}$</td>
<td>-0.094</td>
<td>0.002</td>
<td>0.106</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-0.108</td>
<td>0.008</td>
<td>0.141</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type II</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>SD</td>
<td>RMSE</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>-0.033</td>
<td>0.004</td>
<td>0.069</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.029</td>
<td>0.004</td>
<td>0.071</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.017</td>
<td>0.001</td>
<td>0.042</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.050</td>
<td>0.001</td>
<td>0.063</td>
</tr>
</tbody>
</table>

and the regressor is introduced at all frequencies. Consequently, Assumption 2 is now violated and the estimator faces an additional complication. Nonetheless, simulations do not explicitly reproduce these theoretical results. For instance, the long memory parameters are estimated fairly precisely, especially when $m = \lceil n^{0.8} \rceil$.

Concerning $\xi < 0$, we only report the stationary cointegration case with $\rho = 0$ because the results are very similar as we can see in the lower part of the Table 3. The simulations based on the type I representation are reported in the upper part of the Table 3. We limit our investigation to this case because the computations are highly time consuming but we support that these results are sufficiently informative. We can observe that the differences with the type II representation are sometimes substantial. For instance, whatever the sample, the bias slightly differs when $\delta_1 = 0$. In some cases, the variance increases but the divergences disappear when $\delta_1 = 3$.

Tables 7, 8, 9 and 10 report the results for strong and weak fractional cointegration cases with either $\rho = 0$ or $\rho \neq 0$. In all cases, there is weak evidence of consistency given that RMSEs slowly decrease when the sample size increases. This result suggests that practitioners should devote a particular attention to the stationarity or non-stationarity of the data. In the latter case, they should use the estimator of Hualde (2014).
6. Empirical illustration

In this section we aim to test whether the no-arbitrage condition holds on the crude oil market. A naive approach would be to investigate the presence of cointegration between spot and future prices. But as argued by Brenner and Kroner (1995), they should not be cointegrated because commodity markets are subject to the so-called convenience yield, that is probably not a short memory process. Consequently the long-run equation is contaminated by an additive persistent component. However, as demonstrated by Liu and Tang (2010), the no-arbitrage condition can exist on commodity market if the convenience yield is non-negative and the no-arbitrage issue remains unsolved. Recently, Rossi and Santucci de Magistris (2013), have strengthened that the no-arbitrage condition between spot and future asset prices implies an analogous condition on their underlying volatilities. This original approach is convenient because it can be adapted to commodity markets under mild conditions on the volatility of the convenience yield. Accordingly, we aim to test for the no-arbitrage condition between the crude oil spot and futures volatilities rather than spot and future prices.

The no-arbitrage condition states that on financial markets, risk-free arbitrage opportunities cannot arise. Paradoxically, Grossman and Stiglitz (1980) emphasized that arbitrage opportunities, even infrequent, are necessary to make the market sufficiently incentive for market participants. Investigating the no-arbitrage condition is also interesting because it is closely related to the efficient market hypothesis issue. Denoting $F_{t+k|t}$ and $S_t$ the futures and the spot prices respectively, the no-arbitrage condition implies $F_{t+k|t} = S_t \cdot e^{b_t + u_t}$, with $r_{t+k|t}$ the return of a risk-free asset that expires at time $t+k$ and $e^{b_t+k|t}$ the cost of carry premium. Rossi and Santucci de Magistris (2013) show that using the rescaled daily range, $\sigma_t, X = (log 2)^{-1/2}(max_{\tau} X_{\tau} - min_{\tau} X_{\tau})$, for $X$ any price, the no-arbitrage condition is directly related with the second moments of the spot and futures prices by the following equation,

$$\sigma_{t,F} = \sigma_{t,S} + \beta_t + u_t, \quad \beta_t = (log 2)^{-1/2}(r_{t+k|t} - c_{t+k|t}), \quad t-1 < \tau \leq t$$

where $r_{t+k|t}$ and $r_{t+k|t}$ stand for the risk-free rate in correspondence, respectively, of the highest and lowest log-price in a given day. The additional term $u_t$ is justified by the presence of market frictions whereas $b_t$ could be omitted because the risk-free asset intraday variations are negligible.

Unfortunately, Equation (18) is not correct with respect to our framework because we focus on commodity markets whereas Rossi and Santucci de Magistris (2013) focus on stock index futures and thus neglect the so-called convenience yield. The convenience yield is defined as “the flow of benefit of immediate ownership of a physical commodity” (see Liu and Tang 2011). Reformulating the no-arbitrage condition we obtain

$$F_{t+k|t} = S_t \cdot e^{b_t+k|t} - c_{t+k|t}$$

where $c_{t+k|t}$ represents the net convenience yield that is the difference between the convenience yield
and the cost of storage applicable from time $t$ to time $t + k$. Following the same reasoning as Rossi and Santucci de Magistris (2013), our modified version of Equation (18) should hold if the intraday volatility of the convenience yield is negligible. However, the convenience yield is unobservable and this assumption requires some explanations. Implications of this hypothesis are twofold: (i) it ensures that Equation (18) is not contaminated by the presence of an additional volatility component; (ii) it reduces the probability for the convenience yield to be negative. Indeed, in the latter case, the futures prices would be too high relative to the spot price, thereby offering an arbitrage opportunity. In financial literature, the net convenience yield is often modeled as an Ornstein-Uhlenbeck process. In such a representation the convenience yield volatility is generally high and negative values may frequently occur. Nonetheless, in a recent paper, Liu and Tang (2010) consider Cox-Ingersoll-Ross representation of the convenience yield and demonstrate that its volatility is no necessarily high. Accordingly, in line with their results, we will assume that the intraday volatility of the convenience yield is sufficiently low to be neglected in Equation (18) and to avoid frequent arbitrage-free opportunities.

In the following, we conjecture that when substituting the daily range by the daily squared returns, the relation in Equation (18) remains valid.\footnote{We employ this rough approximation because we do not have the data but we planned to solve this issue before submitting the paper.} Accordingly, we can test whether the no-arbitrage condition holds, by estimating the Equation (18). Nonetheless, the persistent nature of the volatility reveals that long run components drive the underlying processes. In such a case, testing for the presence of cointegration is useful to guard against the risk of spurious regression but also because the dynamics of $\sigma_{t,F}$ and $\sigma_{t,S}$ are likely to slightly diverge in short run. Furthermore, one would expect that the more the maturity of the futures is far away, the more the volatilities are likely to drift far apart. Some evidences of this mechanism are provided by Caporale et al. (2014) on the spot and futures prices.

We focus on WTI crude oil spot and futures prices traded in NYMEX. Our data set runs from January 2, 1996 to December 16, 2013 for a total of $n = 4499$ observations.\footnote{All data were collected on the website of the US Energy Information Administration: http://www.eia.gov/petroleum/.} We consider four different maturity contracts. The contract \textit{Futures 1} specifies the earliest delivery date. It expires on the third business day prior to the 25th calendar day of the month preceding the delivery month. If the 25th calendar day of the month is a non-business day, trading ceases on the third business day prior to the business day preceding the 25th calendar day. The contracts \textit{Futures 2-4} represent the successive delivery months following the contract \textit{Futures 1}. Graphics 2 and 3 represents the daily log squared returns of the spot prices and the four futures.

To test for the presence of stationary (unbalanced) cointegration we apply a rigorous methodology. First of all, we estimate the long memory parameters of each volatility series. Because their is no particular reasons for the volatility to be stationary, we use the two-step exact local Whittle (2S-ELW) estimator of Shimotsu (2010) which is robust to the presence of unknown mean and polynomial time trend. The results are reported in Table 4 for several bandwidths. Clearly, they are fairly homogeneous for a given
bandwidth but heterogeneous across the bandwidths. Such variability is likely to appear in presence of level shifts or short run perturbations. Accordingly, in the following we shall apply the bandwidth
requirement, \( m = o(n^{2\delta/(1+2\delta)}) \), suggested by Frederiksen et al. (2012) to prevent the presence of a neglected additive noise term in the volatility proxy. More precisely, we shall consider two bandwidths: \( m = [n^k] \) for \( k = \{0.4, 0.5\} \).

\[
m = o\left(\frac{n^{2\delta}}{1 + 2\delta}\right), \quad \text{suggested by Frederiksen et al. (2012) to prevent the presence of a neglected additive noise term in the volatility proxy. More precisely, we shall consider two bandwidths:}
\[
m = \lfloor n^k \rfloor \quad \text{for} \quad k = \{0.4, 0.5\}.
\]

### Table 5: W-test for \( m = \lfloor n^k \rfloor \) and \( \epsilon = \{0.02, 0.05\} \)

<table>
<thead>
<tr>
<th>( k )</th>
<th>Spot Prices</th>
<th>Futures 1</th>
<th>Futures 2</th>
<th>Futures 3</th>
<th>Futures 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \epsilon = 0.02 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.6261</td>
<td>0.6832</td>
<td>0.6664</td>
<td>0.6642</td>
<td>0.6642</td>
</tr>
<tr>
<td>0.5</td>
<td>0.8726</td>
<td>1.2638</td>
<td>0.8522</td>
<td>0.8303</td>
<td>0.7964</td>
</tr>
<tr>
<td>( \epsilon = 0.05 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.6261</td>
<td>0.6832</td>
<td>0.6664</td>
<td>0.6642</td>
<td>0.6642</td>
</tr>
<tr>
<td>0.5</td>
<td>0.8726</td>
<td>1.2638</td>
<td>0.8522</td>
<td>0.8303</td>
<td>0.7964</td>
</tr>
</tbody>
</table>

Importantly, we also apply the procedure of Qu (2011) to test the null hypothesis that the volatility estimate is a stationary long memory process against the alternative of a process contaminated by level shifts or a smoothly varying trend. The so-called W-test statistic of Qu (2011) depends on a trimming parameter \( \epsilon \) that we set to either 0.02 or 0.05 and for which there are two specific asymptotic critical value at a threshold of 10%: 1.118 and 1.022. We compute the test using the Gaussian semi-parametric estimator of Robinson (1995a). The results are reported in Table 5 and highlight that in all cases we do not reject the null of a stationary long memory process except for the volatility of the Futures 1 for which it is difficult to conclude.

### Table 6: Unbalance stationary cointegration analysis

<table>
<thead>
<tr>
<th>( k )</th>
<th>Futures 1</th>
<th>Futures 2</th>
<th>Futures 3</th>
<th>Futures 4</th>
<th>Futures 1</th>
<th>Futures 2</th>
<th>Futures 3</th>
<th>Futures 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\delta}_0 )</td>
<td>3.98599</td>
<td>3.87855</td>
<td>3.36281</td>
<td>2.4622</td>
<td>1.22994</td>
<td>1.66711</td>
<td>1.06878</td>
<td>0.287712</td>
</tr>
<tr>
<td>( \hat{\delta}_{0.45} )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \hat{\delta}_{0.35} )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( \hat{\delta}_{0.25} )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

We turn now to the core of our analysis that is the (non) equality of the integration orders. To investigate whether or not the integration orders of the pairwise volatilities are homogeneous, we apply the procedure of Nielsen and Shimotsu (2007). Because the cointegration is not observed, the authors propose to test for \( H_0 : \delta_F = \delta_S \) with \( i = 1, 2, 3, 4 \) that is informative in both cases. However, computing the test statistic, \( T_0 \), requires to estimate the cointegration rank, \( r \) and the authors suggest a model selection procedure based on a tuning parameter \( v_k = m_G^{-k} > 0 \) where \( m_G \) is a specific bandwidth used to estimate
and fixed to \( m_G = \lfloor n_k - 0.05 \rfloor \). Here we consider \( v_k = \{ m_G^{-0.45}, m_G^{-0.35}, m_G^{-0.25} \} \) because the procedure in generally sensitive to the choice of \( \kappa \). Nielsen and Shimotsu (2007) show that \( \hat{T}_0 \xrightarrow{d} \chi^2_1 \) if \( r = 0 \) and \( \hat{T}_0 \xrightarrow{P} 0 \) otherwise. The results are reported in the Table 6 and lead to ambiguous conclusions. Indeed, in several cases we accept the alternative hypothesis at a threshold of 10\% (i.e. 2.71) although the rank estimates are positives in most cases.

In such inconclusive situation we support that unbalanced stationary cointegration is likely to occurs. The estimates \( \hat{\beta}, \hat{\delta}_2, \hat{\delta}_1 \) and \( \hat{\xi} \) are also reported in Table 6. First of all, we observe that all \( \hat{\xi} \) are negative and increases (in absolute value) with the maturity of the futures. According to the 2S-ELW estimates of \( \delta_s \) and \( \delta_{F} \), this is not surprising. Also notice that the cointegration strengths, \( \hat{\nu} \), are clearly non-null and decrease as the maturity of the futures increases. This result is very interesting with respect to hedging strategies and is consistent with the findings of Caporale et al. (2014). Moreover, it reveals that the no-arbitrage hypothesis seems valid for futures contracts maturing in one month whereas the results are more ambiguous for contracts with longer maturities. Indeed, the coefficient \( \beta \) also dramatically decreases as the maturity of the futures increases leading to lower long run hedging ratio for Futures 2-4. A possible explanation comes from the fact that the probability of occurrence of negative convenience yield - and hence the probability arbitrage opportunities - increases with the maturity.

7. Conclusion

In this paper we investigate the estimation of bivariate unbalanced stationary fractional cointegration models. The estimator we propose relies on the local behavior of the spectral density of the system in the vicinity of the origin. It allows for estimating jointly all parameters of interest of the model and notably the unbalanced parameter. We have demonstrated the consistency of the estimator as well as its good finite sample properties estimator by means of Monte Carlo study. Our asymptotic results also suggest that neglecting for the presence of non-null unbalanced parameter might lead to inconsistency. In a short application we investigated the no-arbitrage hypothesis between the volatilities of spot and futures prices. Our results reveal that the apparent unbalance of the integration orders between the daily squared returns of the observable is misleading. An unbalanced stationary cointegration is recovered and the results are consistent with the theory as well as some empirical features found in the literature.

8. Appendix: Proof of theorem 1

**Proof 1.** Let \( \theta \) be the vector of admissible parameter value, \( \theta_0 \) the vector of true parameter value and \( S(\theta) = R_m(\theta) - R_m(\theta_0) \). Then, define the neighborhoods \( \Theta^\delta_\theta = \{ \delta : ||\delta - \delta_0|| < d \} \), \( \Theta^\xi_\theta = \{ \xi : ||\xi - \xi_0|| < e \} \), \( \Theta^\beta_\theta = \{ \beta : ||\lambda_m^{\delta_0} - \delta_0 (\beta - \beta_0)|| < b \} \) and their complements \( \Theta^c_\delta = \Theta_\delta \setminus \Theta^\delta_\theta \), \( \Theta^c_\xi = \Theta_\xi \setminus \Theta^\xi_\theta \) and \( \Theta^c_\beta = \Theta_\beta \setminus \Theta^\beta_\theta \) such that \( \Theta = \Theta_\delta \times \Theta_\xi \times \Theta_\beta \setminus \Xi \). Without loss of generality with respect to Assumption 1 we set

\[
\max_{i} \left( \min ||\delta_i - \delta_{i0}||, ||\xi - \xi_0|| \right) \geq d, \quad \delta \in \Theta^c_\delta, \quad \xi \in \Theta^c_\xi,
\]

(20)
so that $1/2 > d \geq e > 0$. From Robinson (1995a, p. 1634) and by the fact that $\theta_0 \in \Theta_d^0 \times \Theta_e^\xi \times \Theta_\beta^\nu$ it follows

$$
Pr\left(\{\hat{\delta} \in \Theta_d^c \} \cup \{\hat{\xi} \in \Theta_e^c \} \cup \{\hat{\beta} \in \Theta_\beta^c \}\right) =
Pr\left(\inf_{\{\delta \in \Theta_d^c \} \cup \{\xi \in \Theta_e^c \} \cup \{\beta \in \Theta_\beta^c \}} R_m(\theta) \leq \inf_{\{\delta \in \Theta_d^c \} \cup \{\xi \in \Theta_e^c \} \cup \{\beta \in \Theta_\beta^c \}} R_m(\theta) \right)
\leq Pr\left(\inf_{\{\delta \in \Theta_d^c \} \cup \{\xi \in \Theta_e^c \} \cup \{\beta \in \Theta_\beta^c \}} S(\theta) \leq 0 \right).
$$

Accordingly, to prove the Theorem 1, it suffices to show that, as $n \to 0$, $S(\theta)$ is positive and bounded away from 0 uniformly on $\{\hat{\delta} \in \Theta_d^c \} \cup \{\hat{\xi} \in \Theta_e^c \} \cup \{\hat{\beta} \in \Theta_\beta^c \}$ so that

$$
Pr\left(\inf_{\{\delta \in \Theta_d^c \} \cup \{\xi \in \Theta_e^c \} \cup \{\beta \in \Theta_\beta^c \}} S(\theta) \leq 0 \right) \to 0. \quad (21)
$$

Now, introduce $\psi_1 = \delta_1 - \delta_0$ and $\psi_2 = (\delta_2 + \xi) - (\delta_0 + \xi_0)$ and develop $S(\theta)$ as

$$
S(\theta) = \log \det \hat{G}(\theta) - 2(\delta_1 + \delta_2 + \xi)m^{-1}\sum_{j=1}^{m} \log \lambda_j
- \log \det \hat{G}(\theta_0) + 2(\delta_0 + \delta_0 + \xi_0)m^{-1}\sum_{j=1}^{m} \log \lambda_j
= \log \det \hat{G}(\theta) - \log \det \hat{G}(\theta_0) - m^{-1}\sum_{j=1}^{m} 2 \log \lambda_j (\delta_1 + \delta_2 + \xi - \delta_0 - \delta_0 - \xi_0)
= \log \det \hat{G}(\theta) - \log \det \hat{G}(\theta_0) - m^{-1}\sum_{j=1}^{m} 2 \log \lambda_j \sum_{i=1}^{p} \psi_i,
$$

for $p = 2$ so that $\sum_{i=1}^{p} \psi_i = (\psi_1 + \psi_2)$. Then, by the fact that

$$
\sum_{j=1}^{m} \log \lambda_j = \sum_{j=1}^{m} \log (2\pi jn^{-1})
= \sum_{j=1}^{m} \left( \log j + \log (2\pi n^{-1}) \right)
= m \log (2\pi n^{-1}) + \sum_{j=1}^{m} \log j
$$

$$
m^{-1}\sum_{j=1}^{m} \log \lambda_j = \log (2\pi n^{-1}) + m^{-1}\sum_{j=1}^{m} \log j = \log \lambda_m + m^{-1}\sum_{j=1}^{m} \log j - \log m,
$$

18
and rearranging $S(\theta)$, we obtain

$$S(\theta) = \log \det \hat{G}(\theta) - \log \det \hat{G}(\theta_0) + \log \det G_0 - \log \det G$$

$$- 2 \log \lambda_m \sum_{i=1}^{p} \psi_i + 2 \sum_{i=1}^{p} \psi_i \left( \log m - m^{-1} \sum_{j=1}^{m} \log j \right)$$

$$+ \sum_{i=1}^{p} \log(2\psi_i + 1) - \sum_{i=1}^{p} \log(2\psi_i + 1),$$

and finally, $S(\theta) = S_1(\theta) + S_2(\theta) + S_3(\theta)$, where

$$S_1(\theta) = \log \det \hat{G}(\theta) - \log \det G_0 - 2 \log \lambda_m \sum_{i=1}^{p} \psi_i + \sum_{i=1}^{p} \log(2\psi_i + 1)$$

$$S_2(\theta) = \log \det \hat{G}_0 - \log \det \hat{G}(\theta_0)$$

$$S_3(\theta) = 2 \sum_{i=1}^{p} \psi_i \left( \log m - m^{-1} \sum_{j=1}^{m} \log j \right) - \sum_{i=1}^{p} \log(2\psi_i + 1).$$

The way we split $S(\theta)$ has advantage that $S_2(\theta)$ and $S_3(\theta)$ do not depend on $\beta$ so that we can treat them following the methodology of Robinson (1995a). To summarize, we have to demonstrate the boundedness of $S(\theta)$ under

$$\Pr\left( \inf_{\{d \in \Theta_j \cup \tilde{d} \in \Theta_k \cup \beta \in \Theta_p^\prime\}} S(\theta) \leq 0 \right) \leq$$

$$\Pr\left( \inf_{\{d \in \Theta_j \cap \{d \in \Theta_k \cap \beta \in \Theta_p^\prime\}} \left( S_1(\theta) + S_2(\theta) + S_3(\theta) \right) \leq 0 \right) \leq$$

$$\Pr\left( \inf_{\{d \in \Theta_j \cap \{d \in \Theta_k \cap \beta \in \Theta_p^\prime\}} \left( S_1(\theta) + S_2(\theta) + S_3(\theta) \right) \leq 0 \right) \leq$$

$$\Pr\left( \inf_{\{d \in \Theta_j \cap \{d \in \Theta_k \cap \beta \in \Theta_p^\prime\}} \left( S_1(\theta) + S_2(\theta) + S_3(\theta) \right) \leq 0 \right) \leq$$

$$\Pr\left( \inf_{\{d \in \Theta_j \cap \{d \in \Theta_k \cap \beta \in \Theta_p^\prime\}} \left( S_1(\theta) + S_2(\theta) + S_3(\theta) \right) \leq 0 \right) \leq$$

$$\Pr\left( \inf_{\{d \in \Theta_j \cap \{d \in \Theta_k \cap \beta \in \Theta_p^\prime\}} \left( S_1(\theta) + S_2(\theta) + S_3(\theta) \right) \leq 0 \right) \leq$$

$$\Pr\left( \inf_{\{d \in \Theta_j \cap \{d \in \Theta_k \cap \beta \in \Theta_p^\prime\}} \left( S_1(\theta) + S_2(\theta) + S_3(\theta) \right) \leq 0 \right).$$

19
We first consider $S_2(\theta)$ because it does not depend on $\hat{\theta}$ and no uniform bound is needed. Because $|\log (1 + x)| \leq 2|x|$ (see Olver et al. 2010, p. 108), it follows that for $\epsilon \leq 1$

$$\Pr(|S_2(\theta)| \leq \epsilon) = \Pr\left(|\log \det \hat{G}(\theta) - \log \det \hat{G}_0| \leq \epsilon\right) \leq \Pr\left(\left|\frac{\det \hat{G}(\theta_0) - \det \hat{G}_0}{\det \hat{G}_0}\right| \leq \epsilon/2\right).$$

Accordingly, proving that $\det \hat{G}(\theta_0) - \det \hat{G}_0 \xrightarrow{p} 0$ suffices to show that $S_2(\theta)$ is $o_p(1)$ (see Robinson 1995a, p. 1635). To simplify notation, let $F_{jab}^0 = G_{jab}^0 \delta_i - \delta_{ja} \delta_{wb}, f_{jab}^0(\lambda_i) = f_{jab}^0, A_{jab}(e^{\lambda_i}) = A_{jab}$ and $f_{jab}^0 = f_{0jab}(\lambda_i; \theta_{02})$ with $\theta_{02} = (\beta_0, \delta_0')$, $G_{jab}^0$ the $(a,b)$-th element of $G_0$ and $\delta_{0a}$ is the $a$th element of $\theta_{01} = (\delta_{01}, \delta_{02} + \xi_0')$, respectively $\delta_{0b}$. Evaluating Equations (6) and (8) at the true value, we obtain

$$f_j^0 = \begin{pmatrix} I_{xy}(\lambda_i) - 2\beta_0 \lambda_i^2 I_{xx}(\lambda_i) + \beta_0^2 \lambda_i^{2a} I_{xx}(\lambda_i) & I_{xy}(\lambda_i) - \beta_0 \lambda_i^2 I_{xx}(\lambda_i) \\ I_{xy}(\lambda_i) - \beta_0 \lambda_i^2 I_{xx}(\lambda_i) & I_{xx}(\lambda_i) \end{pmatrix},$$

from which we implicitly take the real part. Then,

$$\hat{G}_{ab}(\theta_0) - G_{ab}^0 = \frac{1}{m} \sum_{j=1}^{m} \frac{f_{0jab}^j}{f_{0j}^0} - G_{ab}^0 = \frac{G_{ab}^0}{m} \sum_{j=1}^{m} \left(\frac{f_{0jab}^j}{f_{0j}^0} - 1\right) = \frac{G_{ab}^0}{m} \sum_{j=1}^{m} \left(\frac{f_{0jab}^j}{f_{0j}^0} - 1\right),$$

Recall that under Assumption 3 and 4, $f_2(\lambda_i) = A(e^{\lambda_i})A(e^{\lambda_i})^*/2\pi$ and $f_3(\lambda_i) = I_2/2\pi$. Now, rewrite the latter expression as,

$$\frac{G_{ab}^0}{m} \sum_{j=1}^{m} \left(\frac{f_{0jab}^j}{f_{0j}^0} - 1\right) = S_{21}(\theta) + S_{22}(\theta) + S_{23}(\theta),$$

$$S_{21}(\theta) = \frac{G_{ab}^0}{m} \sum_{j=1}^{m} \left(1 - \frac{f_{0jab}^j}{f_{0j}^0}\right) \frac{f_{0j}^0}{f_{0j}^0},$$

$$S_{22}(\theta) = \frac{G_{ab}^0}{m} \sum_{j=1}^{m} \left(\frac{f_{0jab}^j}{f_{0j}^0} - A_{jab} I_{lb}^0 A_{jab}^0\right),$$

$$S_{23}(\theta) = \frac{G_{ab}^0}{m} \sum_{j=1}^{m} \left(2\pi f_{0jab}^j - 1\right).$$

From the analysis of (Robinson 1995a, p. 1636) and under Assumption 2-5, as $m \to \infty$, $|1 - F_{jab}^0(f_{jab}^0)^{-1}| \leq \eta$, $E|f_{jab}^0(f_{jab}^0)^{-1}| \leq c$ and thus $|S_{21}(\theta)| \leq c\eta$, with $\eta$ and $c$ any arbitrary positive numbers. Then next term to study
is $S_{22}(\theta)$. From Robinson (1995a, p. 1637), it can be demonstrated that

$$E \left| f_{ab}^0 - A_{ab} I_{ac} \bar{A}_{ab} \right| = O \left( f_{ab} \log(j + 1)^{1/2}j^{-1/2} \right),$$

and therefore $|S_{22}(\theta)| = o(1)$ as $m \to \infty$. Finally, as $n \to \infty$ and under Assumption 3,

$$2\pi I_e - I_2 = S_{231}(\theta) + S_{232}(\theta),$$

$$S_{231}(\theta) = n^{-1} \sum_{i=1}^{n} (\varepsilon_i \epsilon_i' - I_2) \xrightarrow{p} 0,$$

$$S_{232}(\theta) = \sum_{s \neq t} n \sum_{i=1}^{n} \cos ((s - t)\lambda_j) \varepsilon_s \epsilon_t = o(1),$$

and $S_{23}(\theta)$ is $o_p(1)$ (see Robinson 1995a, p. 1638). Thereby, as $n \to \infty$, we have proved that

$$\hat{G}_{ab}(\theta_0) - G_{ab}^0 = |S_{21}(\theta)| + |S_{22}(\theta)| + |S_{23}(\theta)|$$

$$= O_p \left( \eta + m^{-1} \sum_{j=1}^{m} \left( \frac{\log j}{j} \right)^{1/2} + o_p(1) \right),$$

and by corollary that $\det \hat{G}(\theta_0) - \det G_0 \xrightarrow{p} 0$. Straightforwardly, $S_2(\theta) = o_p(1)$ follows in Equations (22) to (28).

Now we turn to the analysis of $S_3(\theta)$. First, adding and subtracting $2(\psi_1 + \psi_2)$ to $S_3(\theta)$, we have

$$S_3(\theta) = -2 \sum_{i=1}^{p} \psi_i \left( m^{-1} \sum_{j=1}^{m} \log j - \log m \right) - \sum_{i=1}^{p} \log(2\psi_i + 1) + 2 \sum_{i=1}^{p} \psi_i - 2 \sum_{i=1}^{p} \psi_i$$

Then, rearranging these terms we obtain,

$$S_3(\theta) = -2 \sum_{i=1}^{p} \psi_i \left( m^{-1} \sum_{j=1}^{m} \log j - (\log m - 1) \right)$$

$$+ 2 \sum_{i=1}^{p} \psi_i - \sum_{i=1}^{p} \log(2\psi_i + 1).$$

(29)

From Lemma 2 of Robinson (1995a) we know that

$$m^{-1} \sum_{j=1}^{m} \log j - (\log m - 1) = O(m^{-1} \log m),$$

so that analysis of $S_3(\theta)$ reduces to study the greatest lower bound of (29) which is of the form $f(x) = x - \log(1 +
for \( \psi_1 \) and \((x + y) - \log(x + y + 1)\) for \( \psi_2 \). Because \( \inf_x f(x) \geq x^2 / 6 \) and \( \inf_{x,y} f(x, y) \geq (x^2 + y^2) / 6 \) for \( 0 < |x| < 1, 0 < |y| < 1 \), and from the restriction stated in Equation (20), we can apply the analysis of Nielsen (2007, p. 437) uniformly over \( \{ \delta \in \Theta^c_\delta \} \cup \{ \xi \in \Theta^c_\xi \} \). From Lütkepohl (1996, sec. 8.5.2, p. 111) and by the triangular inequality,

\[
\sqrt{2} \max(|\delta_1 - \delta_{01}|, |\delta_2 - \delta_{02}|) + \sqrt{2} \max(|\xi - \xi_{0}|) \geq \|\Psi\| \geq d + e,
\]

with \( \Psi = (\psi_1, \psi_2)' \). Therefore, the infimum over \( \{ \delta \in \Theta^c_\delta \} \cup \{ \xi \in \Theta^c_\xi \} \cup \{ \beta \in \Theta_\beta \} \) of \( 2 \sum_{i=1}^p \psi_i - \sum_{i=1}^p \log(2\psi_i + 1) \) is no less than

\[
\frac{2(d + e)}{\sqrt{2}} - \log \left( 1 + \frac{2(d + e)}{\sqrt{2}} \right) \geq \frac{2(d^2 + e^2)}{6}.
\]

Then, given that \( f(x) \) has a unique minimum on \((-1, \infty)\) at \( x = 0 \)

\[
\inf_{\{ \delta \in \Theta^c_\delta \} \cup \{ \xi \in \Theta^c_\xi \} \cup \{ \beta \in \Theta_\beta \} } S_3(\theta) \geq \frac{2(d^2 + e^2)}{6} + O(m^{-1} \log m),
\]

\[
\inf_{\{ \delta \in \Theta^c_\delta \} \cup \{ \xi \in \Theta^c_\xi \} \cup \{ \beta \in \Theta_\beta \} } S_3(\theta) = o(1).
\]

The two remaining cases are \( \{ \delta \in \Theta^c_\delta \} \cup \{ \xi \in \Theta^c_\xi \} \cup \{ \beta \in \Theta_\beta \} \) and \( \{ \delta \in \Theta^c_\delta \} \cup \{ \xi \in \Theta^c_\xi \} \cup \{ \beta \in \Theta_\beta \} \). In the former case, \( S_3(\theta) \) is no less than \( 2d^2 / 6 + O(m^{-1} \log m) \) while in the latter, \( S_3(\theta) \) is no less than \( 2e^2 / 6 + O(m^{-1} \log m) \) (see graphic 4).

Figure 4: Plot of \((x + y) - \log(x + y + 1)\) and \((x^2 + y^2) / 6\) (blue and purple curves respectively).

22
Finally we turn to the analysis of $S_1(\theta)$. Rewrite $S_1(\theta)$ as

$$S_1(\theta) = \log \det \hat{G}(\theta) - \log \det G_0 - 2 \log \lambda_m \sum_{i=1}^{p} \psi_i + \sum_{i=1}^{p} \log (2\psi_i + 1)$$

$$= \log \det \hat{G}(\theta) - 2 \log \lambda_m$$

$$\times [\delta_1 - \delta_{01} + (\delta_2 + \xi - \delta_{02} + \xi_0)] - \log \det G_0 + \sum_{i=1}^{p} \log (2\psi_i + 1)$$

$$= \log \det \hat{G}(\theta) - 2 \log \lambda_m \psi_1 - 2 \log \lambda_m \psi_2 - \log \left( \det G_0 (2\psi_1 + 1)^{-1} (2\psi_2 + 1)^{-1} \right)$$

$$= \log \det (V_m \hat{G}(\theta) V_m) - \log \left( \det G_0 (2\psi_1 + 1)^{-1} (2\psi_2 + 1)^{-1} \right),$$

where $\psi_2 = \delta_2 - \delta_{02} = (\delta_2 + \xi) - (\delta_{02} + \xi_0)$, $\det G_0 = G_{11} G_{22} - G_{12}^2$, $V_m = \text{diag}(\lambda_m^{-\psi_1}, \lambda_m^{-\psi_2})$ and

$$\det (V_m \hat{G}(\theta) V_m) = \det \begin{pmatrix} \lambda_m^{-2\psi_1} \hat{G}_{11}(\theta) & \lambda_m^{-\psi_1-\psi_2} \hat{G}_{12}(\theta) \\ \lambda_m^{-\psi_1-\psi_2} \hat{G}_{21}(\theta) & \lambda_m^{-2\psi_2} \hat{G}_{22}(\theta) \end{pmatrix}$$

$$= \lambda_m^{-2\psi_1-2\psi_2} \hat{G}_{11}(\theta) \hat{G}_{22}(\theta) - \lambda_m^{-2\psi_1-2\psi_2} \hat{G}_{21}(\theta) \hat{G}_{12}(\theta)$$

$$= \lambda_m^{-2\psi_1-2\psi_2} (\hat{G}_{11}(\theta) \hat{G}_{22}(\theta) - G_{12}^2(\theta)).$$

Accordingly, analysis of $S_1(\theta)$ reduces to study

$$S_1(\theta) = \lambda_m^{2\psi_1+2\psi_2} \left( \hat{G}_{11}(\theta) \hat{G}_{22}(\theta) - G_{12}^2(\theta) \right) - \left( G_{11} G_{22} - G_{12}^2 \right) (2\psi_1 + 1)^{-1} (2\psi_2 + 1)^{-1}.$$

Because $\hat{G}_{ab}(\theta) = m^{-1} \sum_{j=1}^{m} \lambda_j^{\delta_i+\delta_j} J_{ij}$,

$$S_1(\theta) = S_{11}(\theta) + S_{12}(\theta) + S_{13}(\theta),$$

$$S_{11}(\theta) = \lambda_m^{-2\psi_1-2\psi_2} \left( m^{-1} \sum_{j=1}^{m} \lambda_j^{2\psi_1} J_{11} \times m^{-1} \sum_{j=1}^{m} \lambda_j^{2(\psi_2+\xi)} J_{22} \right),$$

$$S_{12}(\theta) = -\lambda_m^{-2\psi_1-2\psi_2} \left( m^{-1} \sum_{j=1}^{m} \lambda_j^{\delta_i+\delta_j+2\xi} \text{Re}(I_{12}) \right)^2,$$

$$S_{13}(\theta) = - \left( G_{11} G_{22} - G_{12}^2 \right) (1 + 2\psi_1)^{-1} (1 + 2\psi_2)^{-1}.$$

Distinguishing the two summations by indexes $j$ and $k$, multiplying by $G_{11} G_{22} / G_{11} G_{22}$ and $\lambda_j^{-2\psi_1+\lambda_j^{2\psi_2} \lambda_k^{2(\psi_2+\xi)}}$ and

$$\lambda_k^{-2(\delta_{02}+\xi)} \lambda_j^{2(\delta_{02}+\xi_0)},$$

23
and then rearranging $S_{11}(\theta)$ we have

$$S_{11}(\theta) = \lambda_m^{-2\psi_1 - 2\psi_2} \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \lambda_j^{2\delta_k} \lambda_k^{2(\delta_2 + \zeta)} I_{j11} I_{k22} \lambda_j^{-2\delta_0} \lambda_k^{2(\delta_0 + \zeta)} \lambda_k^{-2(\delta_2 + \zeta)}$$

$$= \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \lambda_j^{2\delta_k} \lambda_j^{-2\delta_0} \lambda_k^{2(\delta_2 + \zeta)} \lambda_k^{-2(\delta_0 + \zeta)} \frac{I_{j11} I_{k22}}{\lambda_j^{2\delta_0} \lambda_k^{-2(\delta_0 + \zeta)}}$$

$$= \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{\lambda_j}{\lambda_m} \right)^{2\psi_1} \left( \frac{\lambda_k}{\lambda_m} \right)^{2\psi_2} \frac{I_{j11} I_{k22}}{\lambda_j^{2\delta_0} \lambda_k^{-2(\delta_0 + \zeta)}}$$

$$= \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{j}{m} \right)^{2\psi_1} \left( \frac{k}{m} \right)^{2\psi_2} \frac{I_{j11} I_{k22}}{\lambda_j^{2\delta_0} \lambda_k^{-2(\delta_0 + \zeta)}}. $$

Now, rearranging $S_{12}(\theta)$ we obtain

$$S_{12}(\theta) = -\lambda_m^{-2\psi_1 - 2\psi_2} \left( m^{-1} \sum_{j=1}^{m} \lambda_j^{\delta_1 + \delta_2 + \xi} \text{Re}(I_{j12}) \right)^2$$

$$= -\lambda_m^{-2\psi_1 - 2\psi_2} \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \text{Re}(I_{j12}) \text{Re}(I_{k12}) \lambda_j^{\delta_1 + \delta_2 + \xi} \lambda_k^{\delta_1 + \delta_2 + \xi}$$

$$= -\lambda_m^{-2\psi_1 - 2\psi_2} \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \text{Re}(I_{j12}) \text{Re}(I_{k12}) \lambda_j^{\delta_1 + \delta_2 + \xi} \lambda_k^{\delta_1 + \delta_2 + \xi}$$

$$\times \lambda_j^{\delta_1 + \delta_2 + \xi} \lambda_k^{\delta_1 + \delta_2 + \xi} \lambda_j^{-\delta_1 - \delta_2 - \zeta} \lambda_k^{-\delta_1 - \delta_2 - \zeta} \lambda_j^{\delta_1 + \delta_2 + \xi} \lambda_k^{\delta_1 + \delta_2 + \xi}$$

$$= -\lambda_m^{-2\psi_1 - 2\psi_2} \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \lambda_j^{\delta_1 + \delta_2 + \xi} \lambda_k^{\delta_1 + \delta_2 + \xi} \lambda_j^{-\delta_1 - \delta_2 - \zeta} \lambda_k^{-\delta_1 - \delta_2 - \zeta} \frac{\text{Re}(I_{j12}) \text{Re}(I_{k12})}{\lambda_j^{-\delta_1 - \delta_2 - \zeta} \lambda_k^{-\delta_1 - \delta_2 - \zeta}}$$

$$= -\lambda_m^{-2\psi_1 - 2\psi_2} \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{j}{m} \right)^{2\psi_1 + 2\psi_2} \left( \frac{k}{m} \right)^{2\psi_1 + 2\psi_2} \frac{\text{Re}(I_{j12}) \text{Re}(I_{k12})}{\lambda_j^{-\delta_1 - \delta_2 - \zeta} \lambda_k^{-\delta_1 - \delta_2 - \zeta}}.$$

Then, we follow Nielsen (2007) and correct for the fact that $I_{11}(\lambda)$ and $I_{12}(\lambda)$ are based on estimated cointegration errors. Considering $I_{11}(\lambda) - \tilde{I}_{11}(\lambda)$ and $I_{12}(\lambda) - \tilde{I}_{12}(\lambda)$ we obtain

$$I_{11}(\lambda) = \tilde{I}_{11}(\lambda) - 2\tilde{\beta} \lambda_m^{\delta_0} \text{Re}(I_{12}(\lambda)) + \beta^2 \lambda_m^{2\delta_0} I_{22}(\lambda),$$

$$I_{12}(\lambda) = \tilde{I}_{12}(\lambda) - \tilde{\beta} \lambda_m^{\delta_0} I_{22}(\lambda).$$
with $\tilde{\beta}\lambda_i^m = (\beta - \beta_0)$ and $\nu_0 = \delta_{02} - \delta_{01}$. Substituting in $S_{11}(\theta)$ and $S_{12}(\theta)$ we have

$$S_{11}(\theta) = \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{2\psi_1} \left( \frac{k}{m} \right)^{2\psi_2} \frac{I_{1122}^0}{\lambda_j^{-2\delta_{01}} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$- \frac{2\tilde{\beta}}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{2\psi_1} \left( \frac{k}{m} \right)^{2\psi_2} \frac{\Re(I_{12}^0 I_{22})}{\lambda_j^{-\delta_0} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$+ \frac{\tilde{\beta}^2}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{2\psi_1} \left( \frac{k}{m} \right)^{2\psi_2} \frac{I_{12} I_{22}}{\lambda_j^{-\delta_0} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$S_{12}(\theta) = -\frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{\Re(I_{12}^0 I_{22}^0)}{\lambda_j^{-\delta_0} \lambda_k^{-\delta_0} \lambda_k^{2(\delta_{02} + \delta_0)}}$$

$$+ \frac{2\tilde{\beta}}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12} I_{22}^0}{\lambda_j^{-\delta_0} \lambda_k^{-\delta_0} \lambda_k^{2(\delta_{02} + \delta_0)}}$$

$$- \frac{\tilde{\beta}^2}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{\Re(I_{12}^0 I_{22})}{\lambda_j^{-\delta_0} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$+ \frac{2\tilde{\beta}}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12}^0 I_{22}}{\lambda_j^{-2(\delta_{02} + \delta_0)} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$- \frac{2\tilde{\beta}}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12} I_{22}^0}{\lambda_j^{-2(\delta_{02} + \delta_0)} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$- \frac{\tilde{\beta}^2}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{\Re(I_{12}^0 I_{22})}{\lambda_j^{-2(\delta_{02} + \delta_0)} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$+ \frac{\tilde{\beta}^2}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12} I_{22}^0}{\lambda_j^{-2(\delta_{02} + \delta_0)} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

since $-I_{12} I_{k12} = -I_{12}^0 I_{k12}^0 + 2\beta\lambda_0^m \lambda_i^m \lambda_j^0 I_{12}^0 I_{k12}^0 - \beta^2 \lambda_0^m \lambda_i^m \lambda_j^0 \lambda_k^0 I_{12} I_{k12}^0$. Rearranging this leads to

$$S_{11}(\theta) = \frac{G_{11} G_{22}}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{2\psi_1} \left( \frac{k}{m} \right)^{2\psi_2} \frac{I_{1122}^0}{G_{11} \lambda_j^{-2\delta_{01}} G_{22} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$- \frac{2\tilde{\beta}}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{2\psi_1} \left( \frac{k}{m} \right)^{2\psi_2} \frac{I_{12}^0 I_{22}}{G_{12} \lambda_j^{-\delta_0} G_{22} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$+ \frac{\tilde{\beta}^2}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{2\psi_1} \left( \frac{k}{m} \right)^{2\psi_2} \frac{I_{12} I_{22}}{G_{22} \lambda_j^{-\delta_0} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$S_{12}(\theta) = -\frac{G_{12}}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12}^0 I_{22}^0}{G_{12} \lambda_j^{-\delta_0} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$+ \frac{2\tilde{\beta}}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12}^0 I_{22}}{G_{12} \lambda_j^{-2(\delta_{02} + \delta_0)} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$- \frac{\tilde{\beta}^2}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12} I_{22}}{G_{12} \lambda_j^{-2(\delta_{02} + \delta_0)} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$+ \frac{2\tilde{\beta}}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12}^0 I_{22}^0}{G_{12} \lambda_j^{-\delta_0} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$- \frac{\tilde{\beta}^2}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12} I_{22}}{G_{12} \lambda_j^{-2(\delta_{02} + \delta_0)} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$- \frac{\tilde{\beta}^2}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12}^0 I_{22}^0}{G_{12} \lambda_j^{-\delta_0} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$

$$+ \frac{\tilde{\beta}^2}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \frac{i}{m} \right)^{\psi_1 + \psi_2} \left( \frac{k}{m} \right)^{\psi_1 + \psi_2} \frac{I_{12} I_{22}}{G_{12} \lambda_j^{-2(\delta_{02} + \delta_0)} \lambda_k^{-2(\delta_{02} + \delta_0)}}$$
Now, consider the fact that \( m^{-1} \sum_{j=1}^{m} (j/m)^{\alpha} = (1 + \alpha)^{-1} \) because

\[
m^{-1} \sum_{j=1}^{m} \left( \frac{j}{m} \right)^{\alpha} = \sum_{j=1}^{m} \frac{j}{(j-1)/m} x^{\alpha} dx = \int_{0}^{1} x^{\alpha} dx,
\]

\[
\int_{0}^{1} x^{\alpha} dx = \left[ \frac{1}{1+\alpha} x^{\alpha+1} \right]_{0}^{1} = (1 + \alpha)^{-1},
\]

Moreover, by the analysis of \( \text{(Robinson 1995a, p. 1636-1638)} \) we have

\[
G_{ab} m^{-1} \sum_{j=1}^{m} \left( \frac{\text{Re}(f_{j})}{G_{ab} \hat{\lambda}_{j} - \delta_{a_0} - \delta_{b_0}} - 1 \right) = G_{ab} m^{-1} \left( \sum_{j=1}^{m} \text{Re}(f_{j}) G_{ab}^{-1} \hat{\lambda}_{j} - \delta_{a_0} - \delta_{b_0} - m \right)
\]

\[
= G_{ab} m^{-1} \sum_{j=1}^{m} \text{Re}(f_{j}) G_{ab}^{-1} \hat{\lambda}_{j} - \delta_{a_0} - \delta_{b_0} - 1 = o_p(1),
\]

and thus it follows,

\[
S_{11}(\theta) = G_{11} G_{22} (1 + 2 \psi_1)^{-1} (1 + 2 \psi_2)^{-1} (1 + o_p(1))
- 2 \beta G_{12} G_{22} (1 + 2 \psi_1 - \nu_0)^{-1} (1 + 2 \psi_2)^{-1} (1 + o_p(1))
+ \beta^2 G_{22} (1 + 2 \psi_1 - 2 \nu_0)^{-1} (1 + 2 \psi_2)^{-1} (1 + o_p(1)),
\]

\[
S_{12}(\theta) = -G_{12}^2 (1 + \psi_1 + \psi_2)^{-2} (1 + o_p(1))
+ 2 \beta G_{12} G_{22} (1 + \psi_1 + \psi_2 - \nu_0)^{-1} (1 + \psi_1 + \psi_2)^{-1} (1 + o_p(1))
- \beta^2 G_{22}^2 (1 + \psi_1 + \psi_2 - \nu_0)^{-2} (1 + o_p(1)),
\]

Now we can rewrite \( S_1(\theta) = S_{11}(\theta) + S_{12}(\theta) + S_{13}(\theta) \) as

\[
S_1(\theta) = G_{12}^2 \left( \int_{0}^{1} x^{2 \psi_1} dx \int_{0}^{1} x^{2 \psi_2} dx - \left( \int_{0}^{1} x^{\psi_1 + \psi_2} dx \right)^2 \right)
+ \beta^2 G_{22}^2 \left( \int_{0}^{1} x^{2 \psi_1 - 2 \nu_0} dx \int_{0}^{1} x^{2 \psi_2} dx - \left( \int_{0}^{1} x^{\psi_1 + \psi_2 - \nu_0} dx \right)^2 \right)
+ 2 \beta G_{12} G_{22} \left( \int_{0}^{1} x^{\psi_1 + \psi_2 - \nu_0} dx \int_{0}^{1} x^{\psi_1 + \psi_2} dx - \int_{0}^{1} x^{2 \psi_1 - \nu_0} dx \int_{0}^{1} x^{2 \psi_2} dx \right)
+ G_{11} G_{22} \left( (1 + 2 \psi_1)^{-1} (1 + 2 \psi_2)^{-1} - (1 + 2 \psi_1)^{-1} (1 + 2 \psi_2)^{-1} \right) + o_p(1).
\]

If \( G_{12} = 0 \), by the Cauchy-Schwarz inequality and the fact that \( \nu_0 > 0 \) under cointegration it is straightforward that \( S_1(\theta) \geq o_p(1) \) when \( \{ \delta \in \Theta_{\delta}^c \} \cap \{ \xi \in \Theta_{\xi}^c \} \cap \{ \beta \in \Theta_{\beta}^c \} \), and \( S_1(\theta) \) is bounded away from zero when \( \{ \delta \in \Theta_{\delta}^c \} \cap \{ \xi \in \Theta_{\xi}^c \} \cap \{ \beta \in \Theta_{\beta}^c \} \). A corollary of this finding is that \( S_1(\theta) \) is also bounded when \( \{ \delta \in \Theta_{\delta}^c \} \cap \{ \xi \in \Theta_{\xi}^c \} \cap \{ \beta \in \Theta_{\beta}^c \} \).
\[ \Theta \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \} \text{ and } \{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \}. \]

Similarly, \( S_1(\theta) \geq o_p(1) \) when \( \{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \} \) and \( \{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \}. \) To summarize our results, for \( c > 0 \) an arbitrary small positive number and \( B = 2(d^2 + e^2)/6 + o(1) \), we can rewrite the Equations (22) to (28) as

\[
\Pr \left( \inf_{\{ \hat{\delta} \in \Theta \} \cup \{ \hat{\xi} \in \Theta \} \cup \{ \hat{\beta} \in \Theta \}} S(\theta) \leq 0 \right) \\
\leq \Pr \left( \inf_{\{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \}} \left( S_1(\theta) > c + o_p(1) + [S_3(\theta) \geq B] \right) \leq 0 \right) \\
+ \Pr \left( \inf_{\{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \}} \left( S_1(\theta) \geq o_p(1) + o_p(1) + [S_3(\theta) \geq B] \right) \leq 0 \right) \\
+ \Pr \left( \inf_{\{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \}} \left( S_1(\theta) \geq o_p(1) + o_p(1) + [S_3(\theta) \geq 2d^2/6 + o(1)] \right) \leq 0 \right) \\
+ \Pr \left( \inf_{\{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \}} \left( S_1(\theta) \geq o_p(1) + o_p(1) + [S_3(\theta) \geq 2e^2/6 + o(1)] \right) \leq 0 \right) \\
+ \Pr \left( \inf_{\{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \}} \left( S_1(\theta) > c + o_p(1) + [S_3(\theta) \geq 2d^2/6 + o(1)] \right) \leq 0 \right) \\
+ \Pr \left( \inf_{\{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \}} \left( S_1(\theta) > c + o_p(1) + [S_3(\theta) \geq 2e^2/6 + o(1)] \right) \leq 0 \right) \\
+ \Pr \left( \inf_{\{ \hat{\delta} \in \Theta \} \cap \{ \hat{\xi} \in \Theta \} \cap \{ \hat{\beta} \in \Theta \}} \left( S_1(\theta) > c + o_p(1) + [S_3(\theta) = o(1)] \right) \leq 0 \right),
\]

hence showing the Equation (21) and proving the Theorem 1.

9. Extended Theorem 1

Now we extend the proof of the Theorem 1 to the case where \( G_{12} = G_{21} \neq 0 \). Given that \( G_{12} \) only impact \( S_1(\theta) \) with regard to the positivity of \( S(\theta) \) we only focus on the following expression

\[
S_1(\theta) = G_{12} \left( \int_0^1 x^{\beta_3} dx \int_0^1 x^{2\beta_2} dx - \left( \int_0^1 x^{\beta_1+\beta_2} dx \right)^2 \right) \\
+ \beta^2 G_{22} \left( \int_0^1 x^{2\beta_1-2\beta_2} dx \int_0^1 x^{2\beta_2} dx - \left( \int_0^1 x^{\beta_1+\beta_2} dx \right)^2 \right) + 2\beta G_{12} G_{22} \left( \int_0^1 x^{\beta_1+\beta_2} dx \int_0^1 x^{\beta_1+\beta_2} dx - \int_0^1 x^{\beta_1+\beta_2} dx \int_0^1 x^{\beta_2} dx \right) + o_p(1),
\]

\[
S_1(\theta) = S_{1a}(\theta) + S_{1b}(\theta) + S_{1c}(\theta) + o_p(1).
\]

Now observe that \( S_{1a}(\theta) > 0 \) by the Cauchy-Schwarz inequality. Then, we follow Nielsen (2007) and
consider the fact that \( S_1(\theta) = 0 \) when \( \psi_1 = \psi_2 \). In our unbalanced framework this is a very specific case because \( \psi_2 \) depends on \( \xi \). Also recall that under cointegration, \( \nu_0 > 0 \). Accordingly, for \( \eta > 0 \) and \( |\psi_2 - \psi_1| < \eta \), we have \( S_{1b}(\theta) \geq \beta^2 C \) with \( C > 0 \) and \( |S_{1c}(\theta)| \leq |\beta| \varepsilon \) with \( \varepsilon > 0 \) so that \( -|\beta| \varepsilon \leq S_{1c}(\theta) \leq |\beta| \varepsilon \). Therefore, \( S_1(\theta) \) is no less than \( S_{1b}(\theta) - |\beta| \varepsilon \geq \bar{\beta}^2 C - |\beta| \varepsilon \).

Figure 5: Plot of \( |\bar{\beta}|(|\bar{\beta}|C - \varepsilon) \), \( |\bar{\beta}| \geq 2\varepsilon / C \) and \( |\bar{\beta}| \leq 2\varepsilon / C \) (blue, black and purple curves respectively).

Setting to zero the first derivative of the bound we obtain that \( S_1(\theta) > 0 \) when \( |\bar{\beta}| \geq 2\varepsilon / C \) and no less than \( -\varepsilon^2 / (4C) \) when \( |\bar{\beta}| \leq 2\varepsilon / C \) (see Figure 5). Thus, for \( \eta \) sufficiently small, there exists \( \varepsilon \) such that 22 to 27 tend to 0 when \( |\psi_2 - \psi_1| < \eta \).

Work in Progress

10. Appendix: Auxiliary result

In the following, we extend the Theorem 1 of Robinson and Marinucci (2003) and the Proposition 1 of Nielsen (2005) to the case of unbalanced cointegration. From Equation (8), for \( \delta_1, \delta_2 \) and \( \xi \) consistently pre-estimated, we have

\[
\hat{G}_{ee}(\beta) = \hat{G}_{11}(\beta) = m^{-1} \sum_{j=1}^{m} \left( \lambda_j^{2\delta_1} I_{jj} \right) = m^{-1} \sum_{j=1}^{m} \left( \lambda_j^{2\delta_1} I_{jyy} + \beta^2 \lambda_j^{2\delta_1+2\xi} I_{jxx} - 2\beta \lambda_j^{2\delta_1+\xi} \text{Re}(I_{jxy}) \right),
\]

with \( e_t = y_t - \beta x_t \). Thus, \( R_{m11}(\beta) = \log \hat{G}_{11}(\beta) \) and the derivative with respect to \( \beta \) is \( \partial R_{m11}(\beta) / \partial \beta = 2m^{-1} \sum_{j=1}^{m} \left( \lambda_j^{2\delta_1+\xi} \text{Re}(\beta \lambda_j^{2\delta_1+\xi} I_{jxx} - I_{jxy}) \right) \). Setting this equal to 0 we obtain

\[
\hat{\beta} = \left( m^{-1} \sum_{j=1}^{m} \lambda_j^{2(\delta_1+\xi)} \text{Re}(I_{jxx}) \right)^{-1} \left( m^{-1} \sum_{j=1}^{m} \lambda_j^{2(\delta_1+\xi/2)} \text{Re}(I_{jxy}) \right),
\]
where \( m^{-1}\sum_{j=1}^{m} \lambda_j^{2\delta} \text{Re}(I_{jab}) \) is the \((\delta)\)-weighted periodogram and \( \hat{\beta} \) the narrow-band generalized least squares (NBGLS) estimate of Nielsen (2005).

**Proposition 1.** Under Assumptions 2, 3, 5 and for \( \delta \) satisfying \((2(\delta_2 + \xi) + 2\delta_1 - 1)/4 < \delta \leq \delta_1\),

\[
\hat{\beta} - \beta_0 = \left( m^{-1} \sum_{j=1}^{m} \lambda_j^{2(\delta_1 + \xi)} \text{Re}(I_{jxx}) \right)^{-1} \left( m^{-1} \sum_{j=1}^{m} \lambda_j^{2\delta} \text{Re}(I_{jxe}) \right) = O_p\left( (n/m)^{\delta_1 - \delta_2} \right), \tag{30}
\]

where \( e_t = y_t - \beta x_t \).

**Proof 2.** From the Proposition 1 of Nielsen (2005) we have

\[
m^{-1} \sum_{j=1}^{m} \lambda_j^{2\delta} \text{Re}(I_{jab}) = \int_{0}^{\lambda} \text{Re} \left( f_{ab}(\mu) \right) d\mu \sim \frac{G_{ab}\lambda^{1-\delta_a-\delta_b + 2\delta}}{1-\delta_a - \delta_b + 2\delta} = O_p \left( \lambda_m^{1-\delta_a-\delta_b + 2\delta} \right).
\]

Therefore, by the Cauchy-Schwarz inequality

\[
\left| m^{-1} \sum_{j=1}^{m} \lambda_j^{2\delta} \text{Re}(I_{jxe}) \right| \leq \left( m^{-1} \sum_{j=1}^{m} \lambda_j^{2(\delta_1 + \xi)} \text{Re}(I_{jxx}) \times m^{-1} \sum_{j=1}^{m} \lambda_j^{2\delta_1} \text{Re}(I_{jxe}) \right)^{1/2}
\]

\[
\leq O_p \left( \left( \frac{m}{n} \right)^{1-2(\delta_2 + \xi) + 2(\delta_1 + \xi)} \right)^{1/2} O_p \left( \left( \frac{m}{n} \right)^{1-2\delta_1 + 2\delta_1} \right)^{1/2}
\]

\[
\leq O_p \left( \left( \frac{m}{n} \right)^{1-\delta_2 + \delta_1} \right).
\]

Finally, the proof is completed by substituting this result in Equation (30):

\[
\hat{\beta} - \beta_0 = O_p \left( \left( \frac{m}{n} \right)^{-1+2(\delta_2 + \xi) - 2(\delta_1 + \xi)} \right) O_p \left( \left( \frac{m}{n} \right)^{1-\delta_2 + \delta_1} \right)
\]

\[
= O_p \left( \left( \frac{m}{n} \right)^{\delta_2 - \delta_1} \right).
\]

11. Appendix: Simulation results

As specified in the simulation results section, there is weak evidence of consistency when the parameters lie in the non-stationary regions. Indeed, the RMSE decreases very slowly as the sample size increases. Mostly in the strong cointegration case, the bias increases as the sample size increases. Interestingly, results are less impacted by non-stationarity in the weak cointegration case. This suggests that practitioners should devote a particular attention to the stationarity or non-stationarity of the data. In the latter case, they should use the estimator of Hualde (2014).
Table 7: Simulation results for 10000 replications of the strong cointegration model when $\xi = 0.1$ and $\rho = 0$

<table>
<thead>
<tr>
<th>$m = \lfloor n^{0.5} \rfloor$</th>
<th>Bias</th>
<th>SD</th>
<th>RMSE</th>
<th>Bias</th>
<th>SD</th>
<th>RMSE</th>
<th>Bias</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.6$ $0.0$ $\delta_2$</td>
<td>0.024</td>
<td>0.027</td>
<td>0.166</td>
<td>0.032</td>
<td>0.016</td>
<td>0.131</td>
<td>0.032</td>
<td>0.009</td>
<td>0.101</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.169</td>
<td>0.045</td>
<td>0.271</td>
<td>0.197</td>
<td>0.027</td>
<td>0.257</td>
<td>0.234</td>
<td>0.017</td>
<td>0.267</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.002</td>
<td>0.003</td>
<td>0.055</td>
<td>0.001</td>
<td>0.001</td>
<td>0.036</td>
<td>0.001</td>
<td>0.001</td>
<td>0.027</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>0.039</td>
<td>0.013</td>
<td>0.122</td>
<td>0.043</td>
<td>0.010</td>
<td>0.110</td>
<td>0.050</td>
<td>0.008</td>
<td>0.103</td>
</tr>
<tr>
<td>$0.8$ $0.2$ $\delta_2$</td>
<td>0.007</td>
<td>0.029</td>
<td>0.170</td>
<td>0.004</td>
<td>0.018</td>
<td>0.133</td>
<td>0.001</td>
<td>0.011</td>
<td>0.105</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.185</td>
<td>0.042</td>
<td>0.277</td>
<td>0.217</td>
<td>0.028</td>
<td>0.273</td>
<td>0.259</td>
<td>0.017</td>
<td>0.290</td>
</tr>
<tr>
<td>$\xi$</td>
<td>-0.007</td>
<td>0.005</td>
<td>0.071</td>
<td>-0.007</td>
<td>0.002</td>
<td>0.048</td>
<td>-0.006</td>
<td>0.001</td>
<td>0.034</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>-0.002</td>
<td>0.037</td>
<td>0.193</td>
<td>0.005</td>
<td>0.029</td>
<td>0.169</td>
<td>0.010</td>
<td>0.022</td>
<td>0.148</td>
</tr>
<tr>
<td>$0.9$ $0.4$ $\delta_2$</td>
<td>-0.007</td>
<td>0.031</td>
<td>0.176</td>
<td>-0.007</td>
<td>0.021</td>
<td>0.144</td>
<td>-0.006</td>
<td>0.014</td>
<td>0.117</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.129</td>
<td>0.041</td>
<td>0.240</td>
<td>0.151</td>
<td>0.028</td>
<td>0.224</td>
<td>0.178</td>
<td>0.018</td>
<td>0.223</td>
</tr>
<tr>
<td>$\xi$</td>
<td>-0.026</td>
<td>0.008</td>
<td>0.095</td>
<td>-0.024</td>
<td>0.004</td>
<td>0.069</td>
<td>-0.023</td>
<td>0.002</td>
<td>0.052</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>-0.062</td>
<td>0.063</td>
<td>0.259</td>
<td>-0.070</td>
<td>0.051</td>
<td>0.236</td>
<td>-0.067</td>
<td>0.042</td>
<td>0.215</td>
</tr>
</tbody>
</table>

$m = \lfloor n^{0.8} \rfloor$

| $0.6$ $0.0$ $\delta_2$      | -0.048 | 0.004 | 0.080 | -0.052 | 0.002 | 0.071 | -0.063 | 0.001 | 0.072 |
| $\delta_1$                     | 0.085 | 0.005 | 0.113 | 0.118 | 0.004 | 0.133 | 0.154 | 0.003 | 0.163 |
| $\xi$                          | 0.034 | 0.001 | 0.050 | 0.041 | 0.001 | 0.049 | 0.047 | 0.000 | 0.051 |
| $\hat{\beta}$                 | 0.104 | 0.006 | 0.129 | 0.149 | 0.005 | 0.166 | 0.205 | 0.005 | 0.217 |
| $0.8$ $0.2$ $\delta_2$      | -0.051 | 0.005 | 0.087 | -0.063 | 0.003 | 0.085 | -0.085 | 0.002 | 0.096 |
| $\delta_1$                     | 0.092 | 0.006 | 0.122 | 0.133 | 0.005 | 0.151 | 0.180 | 0.005 | 0.192 |
| $\xi$                          | 0.041 | 0.002 | 0.060 | 0.056 | 0.001 | 0.065 | 0.070 | 0.001 | 0.074 |
| $\hat{\beta}$                 | 0.104 | 0.015 | 0.160 | 0.183 | 0.013 | 0.216 | 0.287 | 0.012 | 0.308 |
| $0.9$ $0.4$ $\delta_2$      | -0.057 | 0.006 | 0.095 | -0.066 | 0.005 | 0.095 | -0.086 | 0.004 | 0.105 |
| $\delta_1$                     | 0.067 | 0.006 | 0.104 | 0.089 | 0.005 | 0.115 | 0.120 | 0.005 | 0.139 |
| $\xi$                          | 0.013 | 0.004 | 0.062 | 0.031 | 0.002 | 0.058 | 0.046 | 0.002 | 0.060 |
| $\hat{\beta}$                 | 0.027 | 0.036 | 0.191 | 0.092 | 0.036 | 0.212 | 0.182 | 0.037 | 0.265 |

References


Table 8: Simulation results for 10000 replications of the strong cointegration model when $\xi = 0.1$ and $\rho = 0.4$

<table>
<thead>
<tr>
<th>$m = \lfloor n^{0.5} \rfloor$</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_2$</td>
<td>Bias</td>
<td>SD</td>
<td>RMSE</td>
</tr>
<tr>
<td>$0.6$</td>
<td>0</td>
<td>$\delta_2$</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>0.157</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>$\xi$</td>
<td>0.028</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>0.175</td>
<td>0.019</td>
</tr>
<tr>
<td>$0.8$</td>
<td>0.2</td>
<td>$\delta_2$</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>0.157</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>$\xi$</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>0.075</td>
<td>0.046</td>
</tr>
<tr>
<td>$0.9$</td>
<td>0.4</td>
<td>$\delta_2$</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>0.111</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>$\xi$</td>
<td>-0.014</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>0.032</td>
<td>0.080</td>
</tr>
</tbody>
</table>

$m = \lfloor n^{0.8} \rfloor$

| $\delta_2$ | Bias | SD  | RMSE | Bias | SD  | RMSE | Bias | SD  | RMSE |
|-------------------|------|------|-----|
| $0.6$ | 0 | $\delta_2$ | -0.067 | 0.004 | 0.091 | -0.066 | 0.002 | 0.079 | -0.068 | 0.001 | 0.074 |
| | $\delta_1$ | 0.108 | 0.005 | 0.130 | 0.134 | 0.004 | 0.148 | 0.159 | 0.003 | 0.167 |
| | $\xi$ | 0.086 | 0.001 | 0.094 | 0.090 | 0.001 | 0.094 | 0.089 | 0.001 | 0.092 |
| | $\beta$ | 0.295 | 0.007 | 0.306 | 0.361 | 0.006 | 0.369 | 0.433 | 0.005 | 0.439 |
| $0.8$ | 0.2 | $\delta_2$ | -0.068 | 0.005 | 0.098 | -0.077 | 0.003 | 0.094 | -0.095 | 0.002 | 0.102 |
| | $\delta_1$ | 0.116 | 0.007 | 0.143 | 0.156 | 0.005 | 0.172 | 0.200 | 0.005 | 0.212 |
| | $\xi$ | 0.082 | 0.002 | 0.094 | 0.096 | 0.001 | 0.101 | 0.108 | 0.001 | 0.111 |
| | $\beta$ | 0.255 | 0.018 | 0.289 | 0.367 | 0.016 | 0.388 | 0.506 | 0.015 | 0.520 |
| $0.9$ | 0.4 | $\delta_2$ | -0.077 | 0.006 | 0.108 | -0.084 | 0.004 | 0.106 | -0.102 | 0.003 | 0.116 |
| | $\delta_1$ | 0.082 | 0.007 | 0.118 | 0.104 | 0.006 | 0.128 | 0.135 | 0.005 | 0.152 |
| | $\xi$ | 0.051 | 0.004 | 0.082 | 0.066 | 0.003 | 0.083 | 0.080 | 0.002 | 0.090 |
| | $\beta$ | 0.191 | 0.050 | 0.294 | 0.279 | 0.050 | 0.358 | 0.399 | 0.051 | 0.459 |

Frederiks, P., Nielsen, F.S., Nielsen, M.Ø., 2012. Local polynomial Whittle estimation of perturbed fractional pro-
Table 9: Simulation results for 10000 replications of the weak cointegration model when $\xi = 0.1$ and $\rho = 0$

<table>
<thead>
<tr>
<th>$m = \lfloor n^{0.5} \rfloor$</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\delta}_2$</td>
<td>Bias</td>
<td>SD</td>
<td>RMSE</td>
</tr>
<tr>
<td>0.6 0.4</td>
<td>$\hat{\delta}_2$</td>
<td>-0.017</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_1$</td>
<td>0.003</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>$\hat{\xi}$</td>
<td>0.013</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>0.094</td>
<td>0.074</td>
</tr>
<tr>
<td>0.8 0.6</td>
<td>$\hat{\delta}_2$</td>
<td>-0.022</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_1$</td>
<td>0.010</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>$\hat{\xi}$</td>
<td>0.020</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>0.139</td>
<td>0.184</td>
</tr>
<tr>
<td>0.9 0.8</td>
<td>$\hat{\delta}_2$</td>
<td>-0.032</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_1$</td>
<td>-0.000</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>$\hat{\xi}$</td>
<td>0.009</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>0.107</td>
<td>0.388</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m = \lfloor n^{0.8} \rfloor$</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\delta}_2$</td>
<td>Bias</td>
<td>SD</td>
<td>RMSE</td>
</tr>
<tr>
<td>0.6 0.4</td>
<td>$\hat{\delta}_2$</td>
<td>-0.052</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_1$</td>
<td>-0.013</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\xi}$</td>
<td>0.035</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>0.179</td>
<td>0.016</td>
</tr>
<tr>
<td>0.8 0.6</td>
<td>$\hat{\delta}_2$</td>
<td>-0.061</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_1$</td>
<td>-0.003</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\xi}$</td>
<td>0.029</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>0.239</td>
<td>0.042</td>
</tr>
<tr>
<td>0.9 0.8</td>
<td>$\hat{\delta}_2$</td>
<td>-0.046</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_1$</td>
<td>-0.014</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\xi}$</td>
<td>-0.014</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>0.058</td>
<td>0.093</td>
</tr>
</tbody>
</table>


Table 10: Simulation results for 10000 replications of the weak cointegration model when $\xi = 0.1$ and $\rho = 0.4$

$m = \lfloor n^{0.5} \rfloor$

<table>
<thead>
<tr>
<th>$\delta_2$</th>
<th>$\delta_1$</th>
<th>$\delta$</th>
<th>Bias</th>
<th>SD</th>
<th>RMSE</th>
<th>Bias</th>
<th>SD</th>
<th>RMSE</th>
<th>Bias</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6 0.4</td>
<td>$\delta_2$</td>
<td>-0.007</td>
<td>0.034</td>
<td>0.185</td>
<td>-0.009</td>
<td>0.022</td>
<td>0.147</td>
<td>0.014</td>
<td>0.013</td>
<td>0.114</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>-0.003</td>
<td>0.032</td>
<td>0.179</td>
<td>-0.011</td>
<td>0.020</td>
<td>0.141</td>
<td>-0.011</td>
<td>0.011</td>
<td>0.107</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\xi$</td>
<td>0.030</td>
<td>0.111</td>
<td>0.111</td>
<td>0.022</td>
<td>0.005</td>
<td>0.076</td>
<td>0.019</td>
<td>0.002</td>
<td>0.053</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>0.413</td>
<td>0.091</td>
<td>0.511</td>
<td>0.434</td>
<td>0.064</td>
<td>0.502</td>
<td>0.455</td>
<td>0.043</td>
<td>0.500</td>
<td></td>
</tr>
<tr>
<td>0.8 0.6</td>
<td>$\delta_2$</td>
<td>-0.019</td>
<td>0.034</td>
<td>0.184</td>
<td>-0.008</td>
<td>0.022</td>
<td>0.149</td>
<td>-0.001</td>
<td>0.013</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>0.002</td>
<td>0.033</td>
<td>0.182</td>
<td>-0.000</td>
<td>0.021</td>
<td>0.145</td>
<td>-0.005</td>
<td>0.012</td>
<td>0.109</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\xi$</td>
<td>0.029</td>
<td>0.016</td>
<td>0.128</td>
<td>0.024</td>
<td>0.008</td>
<td>0.094</td>
<td>0.021</td>
<td>0.004</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>0.407</td>
<td>0.221</td>
<td>0.622</td>
<td>0.421</td>
<td>0.158</td>
<td>0.579</td>
<td>0.435</td>
<td>0.115</td>
<td>0.552</td>
<td></td>
</tr>
<tr>
<td>0.9 0.8</td>
<td>$\delta_2$</td>
<td>-0.037</td>
<td>0.039</td>
<td>0.200</td>
<td>-0.025</td>
<td>0.024</td>
<td>0.158</td>
<td>-0.015</td>
<td>0.017</td>
<td>0.131</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>-0.001</td>
<td>0.032</td>
<td>0.180</td>
<td>-0.007</td>
<td>0.021</td>
<td>0.143</td>
<td>-0.005</td>
<td>0.013</td>
<td>0.116</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\xi$</td>
<td>0.012</td>
<td>0.027</td>
<td>0.166</td>
<td>0.005</td>
<td>0.015</td>
<td>0.123</td>
<td>0.000</td>
<td>0.009</td>
<td>0.094</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>0.374</td>
<td>0.454</td>
<td>0.771</td>
<td>0.350</td>
<td>0.325</td>
<td>0.669</td>
<td>0.336</td>
<td>0.225</td>
<td>0.581</td>
<td></td>
</tr>
</tbody>
</table>

$m = \lfloor n^{0.8} \rfloor$

<table>
<thead>
<tr>
<th>$\delta_2$</th>
<th>$\delta_1$</th>
<th>$\delta$</th>
<th>Bias</th>
<th>SD</th>
<th>RMSE</th>
<th>Bias</th>
<th>SD</th>
<th>RMSE</th>
<th>Bias</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6 0.4</td>
<td>$\delta_2$</td>
<td>-0.064</td>
<td>0.004</td>
<td>0.088</td>
<td>-0.057</td>
<td>0.002</td>
<td>0.071</td>
<td>-0.053</td>
<td>0.001</td>
<td>0.061</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>-0.007</td>
<td>0.003</td>
<td>0.056</td>
<td>-0.005</td>
<td>0.002</td>
<td>0.041</td>
<td>-0.004</td>
<td>0.001</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\xi$</td>
<td>0.065</td>
<td>0.002</td>
<td>0.081</td>
<td>0.065</td>
<td>0.001</td>
<td>0.073</td>
<td>0.061</td>
<td>0.001</td>
<td>0.067</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>0.502</td>
<td>0.017</td>
<td>0.518</td>
<td>0.525</td>
<td>0.012</td>
<td>0.536</td>
<td>0.544</td>
<td>0.010</td>
<td>0.552</td>
<td></td>
</tr>
<tr>
<td>0.8 0.6</td>
<td>$\delta_2$</td>
<td>-0.071</td>
<td>0.004</td>
<td>0.094</td>
<td>-0.065</td>
<td>0.002</td>
<td>0.079</td>
<td>-0.064</td>
<td>0.001</td>
<td>0.073</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>-0.004</td>
<td>0.004</td>
<td>0.059</td>
<td>0.003</td>
<td>0.002</td>
<td>0.045</td>
<td>0.007</td>
<td>0.001</td>
<td>0.038</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\xi$</td>
<td>0.059</td>
<td>0.003</td>
<td>0.079</td>
<td>0.064</td>
<td>0.001</td>
<td>0.074</td>
<td>0.068</td>
<td>0.001</td>
<td>0.074</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>0.523</td>
<td>0.047</td>
<td>0.567</td>
<td>0.569</td>
<td>0.036</td>
<td>0.599</td>
<td>0.623</td>
<td>0.030</td>
<td>0.646</td>
<td></td>
</tr>
<tr>
<td>0.9 0.8</td>
<td>$\delta_2$</td>
<td>-0.053</td>
<td>0.005</td>
<td>0.089</td>
<td>-0.042</td>
<td>0.003</td>
<td>0.068</td>
<td>-0.035</td>
<td>0.002</td>
<td>0.053</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta_1$</td>
<td>-0.015</td>
<td>0.004</td>
<td>0.066</td>
<td>-0.007</td>
<td>0.003</td>
<td>0.051</td>
<td>-0.003</td>
<td>0.002</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\xi$</td>
<td>-0.001</td>
<td>0.004</td>
<td>0.061</td>
<td>0.002</td>
<td>0.002</td>
<td>0.046</td>
<td>0.006</td>
<td>0.001</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>0.339</td>
<td>0.104</td>
<td>0.468</td>
<td>0.342</td>
<td>0.088</td>
<td>0.452</td>
<td>0.365</td>
<td>0.081</td>
<td>0.462</td>
<td></td>
</tr>
</tbody>
</table>


Nielsen, M.O., Shimotsu, K., 2007. Determining the cointegrating rank in nonstationary fractional systems by the exact local Whittle approach. Journal of Econometrics 141, 574-596.


