Igor Viveiros Melo Souza

Tests for Non-Cointegration based on the Frequency Domain

Belo Horizonte
24 de Outubro de 2014

Igor Viveiros Melo Souza

Tests for Non-Cointegration based on the Frequency Domain

Tese apresentada ao Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, para a obtenção de Título de Doutor em Estatística, na Área de Séries Temporais.

Orientador: Valderio Anselmo Reisen

Co-Orientadora: Glaura da Conceição Franco

Belo Horizonte
2014

Igor Viveiros Melo Souza,
Tests for Non-Cointegration based on the Frequency Domain
58 páginas
Tese (Doutorado) - Instituto de Ciências Exatas da Universidade Federal de Minas Gerais. Departamento de Estatística.

1. Cointegração Fracionária
2. Domínio da Frequência
3. Estimador Semiparamétrico
4. Determinante da Matriz de Densidade Espectral

I. Universidade Federal de Minas Gerais. Instituto de Ciências Exatas. Departamento de Estatística.

Comissão Julgadora:

Prof. Dr. Gÿorgy Terdik
(University of Debrecen)

Prof. Dr. Márton Ispány
(University of Debrecen)

Prof. Dr. Flávio Augusto Zeigelmann
(Universidade Federal do Rio Grande do Sul)

Prof. Dr. Bruno de Paula Rocha
(Universidade Federal de Minas Gerais)

Prof. Dr. Valderio Anselmo Reisen
(Universidade Federal do Espírito Santo)

Profa. Dra. Glaura da Conceição Franco
(Universidade Federal de Minas Gerais)

Para minhas três meninas

Agradecimentos

Dificilmente alguém caminha sozinho. Comigo nunca foi diferente. Sempre tive ao meu lado inúmeras pessoas que me ajudaram em diversas dimensões. Portanto, começo por agradecer à minha esposa e minhas filhas. Giselle sempre esteve ao meu lado me dando todo apoio necessário e os resultados que obtive devo grande parte a ela. Às minhas filhas Eleanor e Sofisa peço desculpas por não ter sido mais presente embora sempre tenha me esforçado para não deixar lacunas. Agradeço aos meus pais, Mariza e Marcone, que nunca mediram esforços para me ajudar e aos meus sogros Marlene e Alfredo bem como minha irmã Janaína e ao Daniel. Alguns amigos também foram importantes nessa trajetória. Em especial meus colegas de doutorado Frank e Ivair. Agradeço à nossa querida Geralda Eusébia, ou simplesmente Dona Geralda, por todo esforço e dedicação com as minhas filhas neste período. Sem a sua ajuda, Gêge, eu também não teria consegido ir muito longe.

Por fim gostaria agradecer imensamente aos meus orientadores Glaura e Valderio que compartilharam comigo seus conhecimentos, experiência e, principalmente, paciência. Tive a sorte de tê-los cruzando meu caminho e tem sido um prazer imensurável trabalhar com ambos. Valderio e Glaura são pessoas gentis que além de aturarem meus devaneios foram meus amigos. Agradeço ao Valderio por ter me recebido inúmeras vezes em Vitória e ter dedicado seu espaço pessoal e de trabalho a mim. Obviamente, quaisquer equívocos cometidos aqui são minha inteira responsabilidade.

Resumo

Esta tese se propõe a estudar a cointegração fracionária no domínio da frequência. Aqui investigam-se as restrições que a ausência ou não de cointegração impõe sobre o determinante da matriz de densidade espectral de um vetor de séries bivariado, integrado de ordem 1, quando avaliado na primeira diferença. Permite-se, aqui, que os erros da relação de cointegração sejam fracionalmente integrados. Neste estudo é mostrado que o determinante da matriz de densidade espectral é uma função potência do parâmetro que mensura a redução na ordem de integração do erro (denotado por b) para um conjunto de frequências de Fourier próximas da origem. A partir disto, duas propostas para a estimação do parâmetro de cointegração b são sugeridas. Testes sob a hipótese nula de não cointegração são derivados a partir dos estimadores apresentados e suas propriedades assintóticas discutidas. Estudos com amostras finitas foram realizados com o objetivo de avaliar o desempenho empírico dos estimadores e dos testes propostos através do calculo do vício, do erro quadrático médio, dos níveis de significância e do poder. Os resultados sugerem que os testes possuem níveis de significância empíricos próximos aos níveis nominais. Além disto, o poder dos testes apresenta um desempenho similar quando comparado com o desempenho de outros testes clássicos na literatura de cointegração.

Palavras-chave: Cointegração Fracionária, Domínio da Frequência, Estimador Semiparamétrico, Determinante da Matriz de Densidade Espectral.

Abstract

This thesis proposes to study the fractional cointegration in the frequency domain. Here is investigated the restrictions that the absence or the presence of cointegration imposes on the determinant of the spectral density matrix of a vector of bivariate series, integrated of order 1, when evaluated at the first difference. The errors of the cointegration relationship are allowed to be fractionally integrated. In this study it is shown that the determinant of the spectral density matrix is a power function of the parameter that measures reduction of the order of integration of the error series (denoted here by b) for a set of Fourier frequencies close to the origin. From this, two proposals for the estimation of the cointegrating parameter b are suggested. Tests under the null hypothesis of non-cointegration are derived from these estimators and their asymptotic properties are discussed. A finite sample investigation was conducted in order to evaluate the empirical performance of the estimators and tests by calculating the bias, the mean square error, the significance levels and the power. The results suggest that tests have empirical significance levels close to nominal levels. Furthermore, the power of the tests shows a similar performance compared with the performance of other classical tests in cointegration literature.

Keywords: Fractional cointegration; Determinant of spectral density matrix, Semiparametric estimator.

Lista de Figuras

Lista de Tabelas

Sumário

Capítulo 1

Introdução

O objetivo deste trabalho é estudar cointegração fracionária sob diferentes contextos e propor métodos de estimação dos parâmetros de interesse e testes de não coitegração baseados no domínio da frequência. O conceito de cointegração, introduzido por Granger (1981), tornou-se uma das técnicas mais populares entre os econometristas uma vez que permite verificar se uma combinação linear de séries não estacionárias, integradas de mesma ordem, produz erros cuja ordem de integração é reduzida.

Neste estudo, investigam-se as restrições que a ausência ou não de cointegração impõe sobre o determinante da matriz de densidade espectral de um vetor de séries. Para a pesquisa proposta são considerados vetores bivariados e integrados de ordem 1 e avaliados na primeira diferença com os erros da relação de cointegração fracionalmente integrados. O ponto fundamental sob o qual este trabalho se baseia é o fato de que o determinante da matriz de densidade espectral é uma função potência do parâmetro que reduz a ordem de integração do erro, denotado por b, para um conjunto de frequências de Fourier próximas da origem.

Nesse contexto, no Capítulo 2 é apresentado o artigo *Tests for non-cointegration based on the frequency domain* que é a parte central desta pesquisa. Com base na teoria do domínio da frequência, as propriedades matemáticas e estatísticas dos processos cointegrados são discutidas. Em adição, duas propostas para a estimação do parâmetro de cointegração b são sugeridas: a primeira, baseada em Geweke and Porter-Hudak (1983), propõe uma regressão do logaritmo do determinante da matriz espectral do processo bivariado em estudo. Como segunda proposta, sugere-se um estimador semi-paramétrico do determinante médio baseado na proposta de Robinson (1994).

O artigo também propõe testes sob a hipótese nula de não cointegração, o quais são derivados à partir dos estimadores sugeridos e as propriedades assintóticas desses testes são derivadas. Estudos com amostras finitas foram realizados com o objetivo de avaliar, empiricamente, o desempenho dos estimadores e dos testes propostos por meio do calculo do vício ,do erro quadrático médio, dos níveis de significância e do poder. Os resultados do poder dos testes evidenciaram um desempenho similar comparado com outros testes clássicos na literatura de cointegração discutidos em Dittmann (2000).

A avaliação empírica extende-se por meio de comparação com a metodologia apresentada em Velasco (2003). Esse autor sugere um método alternativo para estimação do parâmetro b. Vale ressaltar que Velasco (2003) apresenta, diferentemente do estudo proposto, as propriedades assintóticas do estimador sob a hipótese de cointegração. Muito embora o estimador sugerido em Velasco (2003) permita que a ordem de integração do vetor seja superior a 1, os resultados das simulações mostram que o método de Velasco (2003) não é robusto a diferentes parametrizações do vetor cointegração, denotado aqui por β, ao passo que as propostas sugeridas nesta tese mostraram-se robustas à variações em β.

Com o objetivo de ilustrar a aplicação dos testes propostos, o artigo apresenta análise de séries reais. Nesse contexto, a hipótese de não cointegração é testada entre as séries do índice Dow Jones da bolsa de valores de Nova Iorque e o índice Financial Times Stock Exchange 100 da bolsa de Londres. Para tal interesse, foram coletadas observações mensais compreendidas entre janeiro de 1985 e maio de 2014.

O terceiro capítulo sugere a utilização de periodogramas robustos nos teste quando as séries possuem *outliers*. O comportamento dos testes robustos é verificado por meio de ensaios empíricos. Por fim, o Capítulo 4 conclui o trabalho com as sugestões de pesquisas futuras.

Capítulo 2

Tests for non-cointegration based on the frequency domain

Igor Viveiros Melo Souza[a], **Valderio Anselmo Reisen**[b], **Glaura da Conceição Franco**[c]

[a]*DECEG, UFOP and Departamento de Estatística, UFMG*

[b]*Departamento de Estatística, UFES and PPGEA, UFES*

[c]*Departamento de Estatística, UFMG*

Abstract The aim of this paper is to propose methods to test the null hypothesis of non-cointegration in bivariate series based on the determinant of the spectral density matrix for the frequencies close to the origin. Two different statistics are proposed: the first one is based on a regression of logged determinant on a set of logged Fourier frequencies and the second statistic is the semiparametric averaged determinant estimator. In the study, series are assumed to be $I(1)$ and the order of integration of the error series is $I(1-b)$, $b \in [0,1]$, that is, the parameter b determines the reduction in the order of integration of the error series. Besides, the determinant of the spectral density matrix for the first difference series is a power function of b. An advantage of the methods proposed here over the standard methods is that they allow to know the order of integration of the error series without estimating a regression equation. Methods discussed here possess correct size and good power for moderate sample sizes when compared with other proposals.

Keywords: Fractional cointegration; Determinant of spectral density matrix, Semiparametric estimator

2.1 Introduction

To study the relationship among economic variables, the concept of cointegration, introduced by Granger (1981), has been widely employed, mainly due to the spurious regression problem. The basic idea of cointegration consists in the fact that a $h \times 1$ vector series $\mathbf{X_t}$, $t = 1, 2, ...$, where each component is non-stationary, can produce some linear combination of its coordinates that has a lower order of integration. After the seminal work of Granger (1981), several studies about this topic have been developed. In the classic context, the most used tests for cointegration are the Engle and Granger (1987) test (EG), the Phillips and Ouliaris (1988) test and the Johansen (1991) procedure. Besides, tests to verify the presence of a unit root are necessary to use appropriate procedures for modeling the data.

Despite its widespread use, the classical set up of cointegration is lately being considered quite restrictive for many real problems. As an alternative, fractional cointegration has emerged as a more adequate methodology and examples to real problems can be seen in Cheung and Lai (1993), Baillie and Bollerslev (1994), Dittmann (2001), McHale and Peel (2010) and Cuestas et al. (2014). Different approaches have been implemented in the estimation and construction of hypothesis tests concerning fractional processes. See, for example, Robinson (1994), Robinson and Marinucci (2001), Marinucci and Robinson (2001), Robinson (2001), Robinson and Yajima (2002) and Velasco (2003).

The first step in cointegration analysis is to verify the order of integration of series $X_{i,t}$, $t = 1, 2, ...,$, $i = 1, ..., h$, that composes the vector $\mathbf{X_t}$. A series $X_{i,t}$ is said integrated of order d, $d \in \Re$, denoted by $X_{i,t} \sim I(d)$, if d is the minimum number of differences required to obtain a process that admits an Autoregressive Moving Average representation (ARMA). In this context, parameter d measures the memory of the series. Using the ARMA representation, series $X_{i,t}$ can be written as:

$$X_{i,t} = (1 - B)^{-d} e_{i,t} \tag{2.1.1}$$

where $e_{i,t} = \theta_q(B)\phi_p^{-1}(B)u_{i,t}$ with $u_{i,t}$ being a white noise process with zero mean and constant variance σ_u^2, $\theta_q(B)$ and $\phi_p(B)$ are polynomials in B with order q and p, respectively, with all roots outside of the unit circle (see Hosking (1981)). B is the backshift operator, that is, $B^\tau X_{i,t} = X_{i,t-\tau}$ $\forall$ $\tau \in \mathbb{N}$. In this case, the series $X_{i,t}$ is said to be an Autoregressive Fractionally Integrated Moving Average process, denoted by ARFIMA (p, d, q). Different values of d attach different properties to the series $X_{i,t}$. A process with $d \leq -0.5$ is stationary but not invertible. When $d \in (-0.5, 0.5)$, $X_{i,t}$ is both stationary and invertible. When $d \geq 0.5$ the process is non-stationary although for $d \in [0.5, 1)$ it is mean-reverting in the sense that innovations do not have long-run impact on the values of the process. For values of $d \geq 1$, the mean-reversion property is no longer valid (for details see Cheung and Lai (1993)).

If $X_{i,t}$ is a stationary process, it has a spectral density function[1], $f_X(\lambda)$, which can be written as:

$$f_X(\lambda) = f_e(\lambda)\left|1 - e^{-i\lambda}\right|^{-2d} \tag{2.1.2}$$

where $f_e(\lambda)$ is the spectral density of a stationary ARMA process e_t with $\lambda \in [0, 2\pi)$ and $d \in \Re$. When the series $X_{i,t}$ is non-stationary, i.e., $d \geq 0.5$, the function defined in Equation 2.1.2 is usually denoted pseudo-spectral density (see, for example, Velasco (2003)).

A general definition of fractional cointegration was given by Robinson and Marinucci (1998), allowing a different order of integration for $X_{i,t}$, that is, $X_{i,t} \sim I(d_i)$, $d_i > 0$, $\forall\ i$. Therefore, the fractional cointegration for a $h \times 1$ vector $\mathbf{X_t}$ is defined as follows:

Definition 1. *Let* $\mathbf{X}_t$*,* $t = 1, 2, ...,$ *be a* $h \times 1$ *vector series whose* i*-th element* $X_{i,t} \sim I(d_i), d_i > 0, i = 1, ..., h$*.* $\mathbf{X}_t$ *is called fractionally cointegrated, denoted by* $\mathbf{X}_t \sim FCI(d_1, ..., d_h, d_\varepsilon)$*, if there exists a* $h \times 1$ *vector* $\boldsymbol{\beta} \neq 0$ *such that* $\varepsilon_t = \boldsymbol{\beta}^T\mathbf{X}_t \sim I(d_\varepsilon)$*, where* $0 < d_\varepsilon < \min_{1 \leq i \leq h} d_i$*.*

The above definition is valid if and only if $d_i = d_j$ for some $i \neq j, i, j = 1, ..., h$. The vector $\boldsymbol{\beta}$ is called cointegration vector. In the case that $d_1 = ... = d_h = d$, it is usual to write $\mathbf{X}_t \sim CI(d, b)$ where $b = d - d_\varepsilon$. When $b = 0$ the vector $\mathbf{X}_t$ is non-cointegrated. In this sense, parameter b measures the reduction in the order

[1] The spectral density of an stationary process X_t is the Fourier transform of the autocovariance function, $\gamma_X(\tau) = \mathbf{E}\left\{(X_{t+\tau} - \mu_X)(X_t - \mu_X)\right\}$, that is, $f_X(\lambda) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty}\gamma_X(\tau)e^{i\tau\lambda}$.

of integration of the error series ε_t.

To test the null hypothesis of fractional non-cointegration, a general approach is to calculate the order of integration of the residual series $\hat{\varepsilon}_t = \hat{\boldsymbol{\beta}}^T \mathbf{X}_t$, after estimating vector $\boldsymbol{\beta}$ (see Dittmann (2000)).

Various estimators of d_ε can be used in hypothesis tests for fractionally cointegrated processes (see, for example, Dittmann (2000) and Santander et al. (2003)). Another approach can be found in Velasco (2003) who proposed a method to estimate and test the parameter b under the null hypothesis of cointegration.

Thus, the main objectives of this work are to propose new methods for estimating the parameter b and, also, a test of non-cointegration based on the determinant of the spectral density matrix of the vector $(\Delta X_{1,t}, \Delta X_{2,t})$, where Δ is the first difference operator, that is, $\Delta = (1 - B)$. Here, special attention is paid to the case where $d = 1$, although the procedures can also be adapted to other cases such as $d \neq 1$. In this situation an appropriate estimator of d is required.

Some theoretical results are established for the proposed methods and an empirical Monte Carlo study is conduct to evaluate their performance for small sample sizes. In addition, the classical cointegration methods are also considered in empirical studies for comparison purposes.

The paper is structured as follows. In Section 2 some properties of the determinant of the spectral density matrix for cointegrated and non-cointegrated series in the first difference are analysed. Section 3 presents the log determinant regression estimator. In addition, the averaged determinant estimator and its modification are discussed. A Monte Carlo study to analyse the performance of proposals suggested here in terms of bias, size and power is presented in Section

4. This section also compares methods proposed here with residual based tests presented in Dittmann (2000) and to the Log Coherency Regression method in Velasco (2003). Section 5 shows an application of the proposed methodologies to a real time series and Section 6 concludes the work.

2.2 The determinant of the spectral density matrix for a bivariate series

This section presents the properties of the determinant of the spectral density matrix for the vector $(\Delta X_{1,t}, \Delta X_{2,t})$ where both components are $I(1)$ and satisfies the linear relationship $X_{1,t} = \beta X_{2,t} + \varepsilon_t$ for some $\beta \neq 0$. The error term ε_t is assumed to be $I(1-b)$, with $0 \leq b \leq 1$, that is, the order of integration can take noninteger values.

If $(X_{1,t}, X_{2,t})$ is cointegrated, i.e., $b \in (0, 1]$, than the determinant of the spectral density matrix of $(\Delta X_{1,t}, \Delta X_{2,t})$ is a power function of b. Let the observable bivariate time series $(X_{1,t}, X_{2,t})$ be formed by the following system:

$$\begin{aligned} X_{1,t} &= \beta_1 T_t + w_{1,t} \\ X_{2,t} &= \beta_2 T_t + w_{2,t} \end{aligned} \tag{2.2.1}$$

for $t = 1, 2, ...,$ $\beta_1 \neq 0$ and $\beta_2 \neq 0$. The series T_t is a common unobservable stochastic trend such that:

$$T_t = (1-B)^{-1}\eta_t \tag{2.2.2}$$

and the innovations η_t are a stationary ARMA process with zero mean such that

$\sum_{\tau=-\infty}^{\infty} |\gamma_{\eta}(\tau)| < \infty$ where $\gamma_{\eta}(\tau)$ is the autocovariance of order τ. The pair of innovations $(w_{1,t}, w_{2,t})$ follows the processes:

$$\begin{aligned} w_{1,t} &= (1-B)^{-(1-b_1)}\, e_{1,t} \\ w_{2,t} &= (1-B)^{-(1-b_2)}\, e_{2,t} \end{aligned} \tag{2.2.3}$$

where $b_1 \in [0,1]$ and $b_2 \in [0,1]$. The vector $(e_{1,t}, e_{2,t})$ follows a zero mean ARMA process with covariance matrix[2] $\mathbf{\Sigma} = \begin{pmatrix} \sum_{\tau=-\infty}^{\infty} \gamma_{e_1}(\tau) & 0 \\ 0 & \sum_{\tau=-\infty}^{\infty} \gamma_{e_2}(\tau) \end{pmatrix}$, such that $\sum_{\tau=-\infty}^{\infty} |\gamma_{e_1}(\tau)| < \infty$, $\sum_{\tau=-\infty}^{\infty} |\gamma_{e_2}(\tau)| < \infty$ and it is uncorrelated with η_t. The system described in Equation 2.2.1 can be rewritten as follows:

$$X_{1,t} = \beta X_{2,t} + \varepsilon_t \tag{2.2.4}$$

where $\beta = \beta_1/\beta_2$, $\beta_1 \neq 0$, $\beta_2 \neq 0$ and $\varepsilon_t = w_{1,t} - (\beta_1/\beta_2)\, w_{2,t}$ is a non-observable error term such that $\varepsilon_t \sim I(1-b)$ with $b = \min(b_1, b_2)$.

Following Definition 1, the vector $(X_{1,t}, X_{2,t})$ will be non-cointegrated if and only if $b = 0$. Note that in Equation 2.2.4, the input series $X_{2,t}$ and error term ε_t are correlated. To impose orthogonality between $X_{2,t}$ and ε_t it is necessary that $\sigma_{u2}^2 = 0$ which implies that $X_{2,t} = T_t$. The spectral density matrix of the vector

[2]Without loss of generality, $\mathbf{\Sigma}$ is assumed to be diagonal in order to avoid the cross spectrum terms between $e_{1,t}$ and $e_{2,t}$ and to make the calculations easier

$(\Delta X_{1,t}, \Delta X_{2,t})$ can be written as (see Priestley (1981) p.658-659):

$$\mathbf{F}(\lambda) = \sum_{\tau=-\infty}^{\infty} \frac{1}{2\pi} \begin{bmatrix} \mathbf{E}\left\{\Delta X_{1,t+\tau}\Delta X_{1,t}\right\} & \mathbf{E}\left\{\Delta X_{1,t+\tau}\Delta X_{2,t}\right\} \\ \mathbf{E}\left\{\Delta X_{2,t+\tau}\Delta X_{1,t}\right\} & \mathbf{E}\left\{\Delta X_{2,t+\tau}\Delta X_{2,t}\right\} \end{bmatrix} e^{-i\lambda\tau}$$
$$= \begin{bmatrix} f_{\Delta X_1}(\lambda) & f_{\Delta X_1 \Delta X_2}(\lambda) \\ f_{\Delta X_2 \Delta X_1}(\lambda) & f_{\Delta X_2}(\lambda) \end{bmatrix}, \quad (2.2.5)$$

where $f_{\Delta X_1}(\lambda)$ and $f_{\Delta X_2}(\lambda)$ are the spectral densities of $\Delta X_{1,t}$ and $\Delta X_{2,t}$, respectively and $f_{\Delta X_1 \Delta X_2}(\lambda)$ and $f_{\Delta X_2 \Delta X_1}(\lambda)$ are the cross-spectrum between $\Delta X_{1,t}$ and $\Delta X_{2,t}$. The matrix $\mathbf{F}(\lambda)$ is Hermitian which means that $f_{\Delta X_1 \Delta X_2}(\lambda) = \overline{f_{\Delta X_2 \Delta X_1}(\lambda)}$, where the over line means the complex conjugate.

Using the standard spectral properties of multivariate time series, $\mathbf{F}(\lambda)$ can be rewritten as (see Priestley (1981)):

$$\mathbf{F}(\lambda) = \begin{bmatrix} \beta_1^2 f_\eta(\lambda) + \left|1 - e^{-i\lambda}\right|^{2b_1} f_{e_1}(\lambda) & \beta_1\beta_2 f_\eta(\lambda) \\ \beta_2\beta_1 f_\eta(\lambda) & \beta_2^2 f_\eta(\lambda) + \left|1 - e^{-i\lambda}\right|^{2b_2} f_{e_2}(\lambda) \end{bmatrix}. \quad (2.2.6)$$

The determinant of matrix $\mathbf{F}(\lambda)$ is:

$$D(\lambda) = \left|1 - e^{-i\lambda}\right|^{2b_1} \beta_2^2 f_{e_1}(\lambda) f_\eta(\lambda) + \left|1 - e^{-i\lambda}\right|^{2b_2} \beta_1^2 f_{e_2}(\lambda) f_\eta(\lambda) +$$
$$\left|1 - e^{-i\lambda}\right|^{2(b_1+b_2)} f_{e_1}(\lambda) f_{e_2}(\lambda). \quad (2.2.7)$$

Assuming without loss of generality that $b_1 \leq b_2$, which makes $b = b_1$, and using the fact that,

$$\left|1 - e^{-i\lambda}\right|^{2b^*} = (2 - 2\cos\lambda)^{b^*},$$

and:

$$\lim_{\lambda \to 0^+} \frac{(2 - 2\cos\lambda)^{b^*}}{\lambda^{2b^*}} = 1,$$

which means that for $b^* \in \Re$, $|1 - e^{-i\lambda}|^{2b^*} = O(\lambda^{2b^*})$, the determinant $D(\lambda)$ can be rewritten as:

$$D(\lambda) = |1 - e^{-i\lambda}|^{2b} \frac{G(\lambda)}{G(0)} G(0) + O(\lambda^{2b_2}) + O(\lambda^{2(b+b_2)}), \qquad (2.2.8)$$

where $G(\lambda)$ is a bounded function due to the stationarity of the processes $e_{1,t}$, $e_{2,t}$ and η_t such that:

$$\lim_{\lambda \to 0^+} \frac{G(\lambda)}{G(0)} = 1.$$

From this, the determinant $D(\lambda)$ can be computed as:

$$D(\lambda) \sim |1 - e^{-i\lambda}|^{2b} \frac{G(\lambda)}{G(0)} G(0) \ as \ \lambda \longrightarrow 0^+, \qquad (2.2.9)$$

where the symbol " $\sim$ " means that ratio of left and right-hand sides tends to a constant $0 < C < \infty$ as $\lambda \to 0^+$. From the Equation 2.2.9, $D(\lambda)$ depends on the reduction of the order of integration b imposed by cointegration. Similar results are also described by Nielsen (2004). Therefore, if $(X_{1,t}, X_{2,t})$ is cointegrated, that is, $0 < b \leq 1$, $D(\lambda) \to 0$ as $\lambda \longrightarrow 0^+$. It means that $\mathbf{F}(\lambda)$ is a matrix with incomplete rank at $\lambda = 0$ (see Phillips and Ouliaris (1988)). In the case of $b = 0$, that is, $(X_{1,t}, X_{2,t})$ is non-cointegrated, $D(\lambda) \to C$ as $\lambda \longrightarrow 0^+$ and $\mathbf{F}(\lambda)$ has full rank at $\lambda = 0$.

Therefore, new methods to estimate parameter b and test the null hypothesis of non-cointegration are proposed by analysing the slope of the function $D(\lambda)$ in a neighborhood of zero frequency.

2.3 Estimating b

Standard estimation methods for the memory parameter d, well discussed in the literature of long memory processes, can be used as alternative procedures to obtain estimates of b. These procedures are addressed here using the fact that $D(\lambda) \sim O(\lambda^{2b})$. The first proposal is similar to the approach of Geweke and Porter-Hudak (1983), where the logged periodogram is regressed on logged Fourier frequencies. The second one is based on Robinson (1994) semiparametric averaged periodogram estimator of d, where a logged ratio of the periodogram is evaluated in a neighborhood of zero frequency.

2.3.1 The logged determinant regression

Similar to the estimator of d proposed by Geweke and Porter-Hudak (1983) (GPH), an estimate of b can be computed from an approximated regression equation of $\ln D(\lambda) \sim 2\ln|1-e^{-i\lambda}|$ when $\lambda \longrightarrow 0^+$. By taking the log in the Equation 2.2.9 yields:

$$\ln D(\lambda) \sim \ln G(0) + \ln \frac{G(\lambda)}{G(0)} + b \ln|1-e^{-i\lambda}|^2 \; as \; \lambda \longrightarrow 0^+. \tag{2.3.1}$$

For a pair of series $(\Delta X_{1,t}, \Delta X_{2,t})$ with a sample of size n, ie, $t = 1, ..., n$, the first step in order to implement the above regression model is to estimate the spectral density matrix, $\mathbf{F}(\lambda)$, in 2.2.6. Let $\lambda_j = 2\pi j/n$, $j = l, l + (2r+1), l +$

$2(2r+1), ..., m-(2r+1), m$, where $r, l \in \mathbb{N}^*$, with $r < l < m$ and $m < n$. Hence, the estimate of $\mathbf{F}(\lambda_j)$ is given by

$$\widehat{\mathbf{F}}_r(\lambda_j) = \frac{1}{(2r+1)} \begin{bmatrix} \sum_{v=j-r}^{j+r} I_{n,\Delta X_1}(\lambda_v) & \sum_{v=j-r}^{j+r} I_{n,\Delta X_1 \Delta X_2}(\lambda_v) \\ \sum_{v=j-r}^{j+r} I_{n,\Delta X_2 \Delta X_1}(\lambda_v) & \sum_{v=j-r}^{j+r} I_{n,\Delta X_2}(\lambda_v) \end{bmatrix}, \qquad (2.3.2)$$

where each diagonal term of $\widehat{\mathbf{F}}_r(\lambda_j)$ is the average of $2r+1$ distinct periodograms centered at frequency j given by:

$$I_{n,\Delta X_i}(\lambda_j) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} X_{i,t} e^{-i\lambda_j t} \right|^2, \qquad (2.3.3)$$

for $i = 1, 2$. The off-diagonal terms of $\widehat{\mathbf{F}}_r(\lambda_j)$ are also an average of $2r+1$ distinct cross-periodograms centered at frequency j that can be computed by:

$$I_{n,\Delta X_s \Delta X_p}(\lambda_j) = \left(\sum_{t=1}^{n} \Delta X_{s,t} e^{-i\lambda_j t} \sum_{t=1}^{n} \Delta X_{p,t} e^{i\lambda_j t} \right) / 2\pi n \qquad (2.3.4)$$

where $p, s = 1, 2$, $p \neq s$. The natural estimate of $D(\lambda_j)$ in Equation 2.3.1 is the determinant of $\widehat{\mathbf{F}}_r(\lambda_j)$ denoted here by $\widehat{D}_r(\lambda_j)$. Equation 9.5.12 from Priestley (1981), p.697, states that $cov\left[I_{n,\Delta X_{s_1} \Delta X_{p_1}}(\lambda_j), I_{n,\Delta X_{s_2} \Delta X_{p_2}}(\lambda_k)\right] \to 0$ as $n \to \infty$, where $s_1, p_1, s_2, p_2 = 1, 2$, and $\lambda_j = 2\pi j/n$, $\lambda_k = 2\pi k/n$, $j, k = 1, ..., n$ with $j \neq k$. Since the quantities in $\widehat{\mathbf{F}}_r(\lambda_j)$ are calculated with non-overlapping Fourier frequencies, they satisfy the conditions presented by Priestley (1981) in order to be asymptotically uncorrelated and, as a result, $cov\left[D(\lambda_j), D(\lambda_k)\right] \to 0$ as $n \to \infty$. If the process $(\Delta X_{1,t}, \Delta X_{2,t})$ is Gaussian, then $cov\left[D(\lambda_j), D(\lambda_k)\right] = 0$ $\forall n$.

Using the fact that $\ln\left[\frac{\widehat{D}_r(\lambda_j)}{D(\lambda_j)}\right] - \ln \widehat{D}_r(\lambda_j) = -\ln D(\lambda_j)$ and replacing $D(\lambda_j)$ by the approximation in Equation 2.3.1, the following regression equation is obtained:

$$\ln \widehat{D}_r(\lambda_j) = \ln G(0) + \ln \frac{G(\lambda_j)}{G(0)} + c(r) + b \ln |1 - e^{-i\lambda_j}|^2 + \left\{ \ln \left[\frac{\widehat{D}_r(\lambda_j)}{D(\lambda_j)} \right] - c(r) \right\} \tag{2.3.5}$$

where $c(r) = \mathbf{E}\left\{\ln\left[\frac{\widehat{D}_r(\lambda_j)}{D(\lambda_j)}\right]\right\}$.

Therefore, the ordinary least squares estimator of b, $\hat{b}_{LDR}$, is:

$$\hat{b}_{LDR} = \left(\sum_{j=l}^{m} \tilde{Z}_j^{\,2} \right)^{-1} \sum_{j=l}^{m} \tilde{Z}_j (\ln \widehat{D}_r(\lambda_j)), \tag{2.3.6}$$

where $Z_j = \ln(2 - 2\cos\lambda_j)$ and $\tilde{Z}_j = Z_j - \bar{Z}$, $\bar{Z}$ is the mean of Z_j. In order to obtain some asymptotic results of $\hat{b}_{LDR}$, under the null hypothesis of non-cointegration, the following assumptions are introduced:

Assumption 1. *The vector of innovations $(\Delta X_{1,t}, \Delta X_{2,t})$ follows a Gaussian white noise process with zero mean and covariance matrix $\mathbf{\Sigma}$.*

Assumption 2. *Let $m = g(n)$ such that $\frac{g(n)}{n} + \frac{1}{g(n)} + \frac{\ln n}{g(n)} \to 0$ as $n \to \infty$.*

Remark 1. *Under Assumption 1, the spectral density of $(\Delta X_{1,t}, \Delta X_{2,t})$, $\mathbf{F}(\lambda_\mathbf{j})$, is constant across different values of λ_j. In other words, the value of $\mathbf{F}(\lambda_\mathbf{j})$ is independent of j, $j = 1, ..., m$, and, therefore, $D(\lambda_j) = \Lambda$, where Λ is a positive constant. Moreover, if Assumption 1 holds, than the system described by Equation 2.2.1 is necessarily non-cointegrated, that is, $b = b_1 = b_2 = 0$.*

Following Goodman (1963) and under Assumption 1, the distribution of the quantity $\mathscr{D} = \ln\left[4\,(2r+1)^2\,\widehat{D}_r(\lambda_j)/\Lambda\right]$ has the same properties of $ln(\chi^2_{(4r+2)}\chi^2_{(4r)})$, that is, $\mathscr{D} \overset{d}{=} ln(\chi^2_{(4r+2)}\chi^2_{(4r)})$ where $\chi^2_{(4r+2)}$ and $\chi^2_{(4r)}$ are chi-squared random variables with $(4r+2)$ and $4r$ degrees of freedom, respectively. The symbol $\overset{d}{=}$ means equality in distribution. In addition, since the vector $(\Delta X_{1,t}, X_{2,t})$ is a white noise, $G(\lambda)$ in Equation 2.2.8 is constant, that is, $G(\lambda) = G(0)$.

Proposition 1. *Let the bivariate time series $(X_{1,t}, X_{2,t})$ satisfying the Equation 2.2.4. If Assumptions 1 and 2 hold, then:*

1. $\mathbf{E}\left[\hat{b}_{LDR}\right] = 0$;
2. $\mathbf{V}\left[\hat{b}_{LDR}\right] = \frac{\psi^{(1)}(2r+1)+\psi^{(1)}(2r)}{\sum_{j=l}^{m} \tilde{Z}_j^{\,2}}$;
3. $\mathbf{V}\left[\hat{b}_{LDR}\right] \to 0$ *as* $m \to \infty$,

where $\psi^{(1)}(z)$ is the Polygamma function of order 1, that is, $\psi^{(1)}(z) = \frac{d^2 \ln\Gamma(z)}{dz^2}$.

Assumption 1 is quite strong but necessary to understand the behavior of the statistic $\hat{b}_{LDR}$. As pointed out, under the assumption of non-cointegration and for a fixed m and r, the distribution of $\hat{b}_{LDR}$ will be the same of a weighted sum of $\{W_j\}_{j=l}^{m}$ independent random variables where each $W_j \overset{d}{=} ln(\chi^2_{(4r+2)}\chi^2_{(4r)})$ and the weights are $\tilde{Z}_j / \left(\sum_{j=l}^{m} \tilde{Z}_j^{\,2}\right)$. Once $m \to \infty$ the following proposition can be stated:

Proposition 2. *Let Assumptions 1 and 2 hold. Then, for a fixed positive integer r*

$$\mathbf{V}\left[\hat{b}_{LDR}\right]^{-1/2} \hat{b}_{LDR} \overset{d}{\to} N(0,1)$$

as $m \to \infty$, where $\hat{b}_{LDR}$ and $\mathbf{V}\left[\hat{b}_{LDR}\right]$ are given by Equations 2.3.6 and .2.7, respectively.

The proof of Propositions 1 and 2 can be viewed in Appendixes .2 and .3, respectively. Assumption 1 can be relaxed allowing the terms $D(\lambda_j)$ and $\ln G(\lambda_j)$ to vary across different frequencies. In that case, the estimator $\hat{b}_{LDR}$ is still consistent since for sufficiently close frequencies λ_j, λ_k, $j \neq k$, $D(\lambda_j) \approx D(\lambda_k)$ as $n \to \infty$ and $\ln G(\lambda_j) \to \ln G(0)$ as $m \to \infty$.

2.3.2 The Averaged Determinant

Based on the semiparametric averaged periodogram estimator of d proposed by Robinson (1994), an alternative to estimate b can be computed due to the fact that the $D(\lambda)$ in Equation 2.2.9 is a regularly varying function of index $2b$, that is: [3]

$$\lim_{\lambda \to 0^+} \frac{D(q\lambda)}{D(\lambda)} = q^{2b}, \tag{2.3.7}$$

where q is a positive constant. Based on Equation 2.3.7,

$$b \cong (2 \ln q)^{-1} \ln \frac{D(q\lambda)}{D(\lambda)}. \tag{2.3.8}$$

The estimate of b, say $\hat{b}_{AD}$, is computed by replacing $D(\cdot)$ by its estimate and

[3] Following the definition given by Bingham et al. (1987), a measurable function $\mathcal{H} : [a^*, \infty) \to (0, \infty)$, $\forall a^* \in \Re$, is said to be regularly varying of index ϑ, $\vartheta \in \Re$, if $\mathcal{H}$ satisfies:

$$\lim_{y \to \infty} \frac{\mathcal{H}(\alpha y)}{\mathcal{H}(y)} = \alpha^{\vartheta} \; \forall \alpha > 0.$$

In the present case, let $y = 1/\lambda$. Thus, for a positive q:

$$\lim_{\lambda \to 0^+} \frac{D(q\lambda)}{D(\lambda)} = \lim_{y \to \infty} \frac{\mathcal{H}(qy)}{\mathcal{H}(y)} = q^{2b}$$

this is discussed as follows. For a sample of $(\Delta X_{1,t}, \Delta X_{2,t})$, $t = 1, ..., n$, let now the estimate of $\mathbf{F}(\lambda_j)$, $j = 1, ..., m$, be given by:

$$\widehat{\mathscr{F}}(\lambda_j) = \begin{bmatrix} I_{\Delta X_1 \Delta X_1}(\lambda_j) & I_{\Delta X_1 \Delta X_2}(\lambda_j) \\ I_{\Delta X_2 \Delta X_1}(\lambda_j) & I_{\Delta X_2 \Delta X_2}(\lambda_j) \end{bmatrix}. \tag{2.3.9}$$

The estimate of $D(\lambda)$ and $D(q\lambda)$ are then obtained as follows:

$$\widehat{D}(\lambda_m) = \left\| \sum_{j=1}^{m} \widehat{\mathscr{F}}(\lambda_j) \right\| \quad \text{and} \quad \widehat{D}(q\lambda_m) = \left\| \sum_{j=1}^{m} \widehat{\mathscr{F}}(q\lambda_j) \right\|, \tag{2.3.10}$$

where $\|\mathbf{A}\|$ denotes the determinant of the matrix $\mathbf{A}$ and m satisfies Assumption 2. Therefore, $\widehat{D}(q\lambda_m) \to \widehat{D}(0)$ and $\widehat{D}(\lambda_m) \to \widehat{D}(0)$. The parameter b is estimated by:

$$\widehat{b}_{AD} = (2 \ln q)^{-1} \ln \frac{\widehat{D}(q\lambda_m)}{\widehat{D}(\lambda_m)}. \tag{2.3.11}$$

The statistic in 2.3.11 will be called the Average Determinant (AD) estimator. Under Assumption 1, that is, $(\Delta X_{1,t}, \Delta X_{2,t})$ follows a Gaussian white noise process, $\widehat{b}_{AD}$ can be rewritten as:

$$\widehat{b}_{AD} = \left\{ \ln \left[\widehat{D}(q\lambda_m)/\Lambda \right] - \ln \left[\widehat{D}(\lambda_m)/\Lambda \right] \right\} (2 \ln q)^{-1}. \tag{2.3.12}$$

Since λ_j is the Fourier frequency, the set of variables $\widehat{\mathscr{F}}(\lambda_j)$ are independently distributed and each is asymptotically distributed as a 2×2 complex Whishart matrix, that is, $\widehat{\mathscr{F}}(\lambda_j) \sim \mathbf{W}_2{}^c(1, \mathbf{f}(\lambda))$ (See Brillinger (1981) pp. 305).

In the case where q is not a positive integer number, the quantity $2\pi jq/n$ is no longer a Fourier frequency and, in order to guarantee asymptotic independence of

$\widehat{\mathscr{F}}(q\lambda_j)$, the frequencies λ_j should be chosen such that λ_j and λ_k, $j,k = 1,...,m$, $j \neq k$, are spaced sufficiently apart. In particular, $|\lambda_j \pm \lambda_k| \gg 2\pi/n$, which is equivalent to $|j \pm k| \gg 1/q$ (see Priestley (1981), pp. 405). This condition is easily achieved, for example, if $q \in (0.5, 1)$ and $|j \pm k| \geq 2$, that is, j and k are chosen from a set of odd or even numbers.

Note that the asymptotical independence between $\widehat{D}(q\lambda)$ and $\widehat{D}(\lambda)$ is not always guaranteed since for some set of frequencies $(q\lambda_j, \lambda_k)$, $|q\lambda_j \pm \lambda_k| < 2\pi/n$. In order to solve this problem the frequencies can be trimmed out, although this can lead to very poor estimates in practical situations, as the estimates will be calculated with a reduced number of frequencies.

Under Assumptions 1 and 2, $\hat{b}_{AD}$ is an unbiased estimator of b and consistent. This can be summarized in the following proposition:

Proposition 3. *Let the bivariate time series $(X_{1,t}, X_{2,t})$ satisfying the Equation 2.2.4. If Assumptions 1 and 2 hold, then:*

1. $\mathbf{E}\left[\hat{b}_{AD}\right] = 0$;
2. $\mathbf{V}\left[\hat{b}_{AD}\right] = \left(\psi^{(1)}(m-1) + \psi^{(1)}(m)\right)(1-\rho) \,/\, \left\{2\left(\ln q\right)^2\right\}$;
3. $\mathbf{V}\left[\hat{b}_{AD}\right] \to 0$ *as* $m \to \infty$,

where ρ is the correlation coefficient between $\widehat{D}(q\lambda)$ and $\widehat{D}(\lambda)$ which has no closed form if no frequencies are trimmed out. The proof of Proposition 3 is in Appendix .4.

Corollary 1 of Cai et al. (2013) shows that $\frac{\ln\left(\widehat{D}(\lambda_m)/\Lambda\right) - 3/\lceil m/2\rceil}{2/\sqrt{\lceil m/2\rceil}} \xrightarrow{d} N(0,1)$. Based on this, the above estimator can be written as a sum of two asymptotic correlated Gaussian processes. Then, heuristically, it should converge to a Gaussian random variable.

As an alternative to AD estimator,one should consider all frequencies in the denominator of $\frac{\widehat{D}(q\lambda_m)}{\widehat{D}(\lambda_m)}$ which has variance reduction compared with $\hat{b}_{AD}$. This alternative method, denoted by Modified Averaged Determinant (MAD), is discussed now. In this case, the numerator will take into account $\lceil m/2 \rceil$ distinct frequencies while the denominator will take into account $[m]$ distinct frequencies. Let m_1 and m_2 be the number of frequencies included in the numerator and denominator, respectively. Using Equation .2.2 in Appendix .2 one can see that, under the hypothesis of non-cointegration,

$$\mathbf{E}\left[\hat{b}_{AD}\right] = 2(\ln m_2 - \ln m_1) + \psi^{(0)}(m_1) + \psi^{(0)}(m_1 - 1) - \psi^{(0)}(m_2) - \psi^{(0)}(m_2 - 1), \tag{2.3.13}$$

where $\psi^{(0)}(z)$ is the digamma function: $\psi^{(0)}(z) = \frac{d \ln \Gamma(z)}{dz}$. Since $\widehat{b}_{AD}$ will no longer be centered at $b = 0$, a bias correction must be considered. In this sense, the MAD estimator is defined by:

$$\begin{aligned} \widehat{b}_{MAD} = \frac{\ln \frac{\widehat{D}(q\lambda_{m_1})}{\widehat{D}(\lambda_{m_2})}}{2 \ln q} &- [2(\ln m_2 - \ln m_1) + (\psi^{(0)}(m_1) + \psi^{(0)}(m_1 - 1)) \\ &- (\psi^{(0)}(m_2) + \psi^{(0)}(m_2 - 1))]/(2 \ln q). \end{aligned} \tag{2.3.14}$$

The variance of $\widehat{b}_{MAD}$ will be:

$$\mathbf{V}\left\{\widehat{b}_{MAD}\right\} = \left\{\left[\psi^{(1)}(m_1-1)+\psi^{(1)}(m_1)\right]+\left[\psi^{(1)}(m_2-1)+\psi^{(1)}(m_2)\right] - 2\rho\sqrt{\psi^{*}\left(m_1,m_2\right)}\right\} / \left\{4\left(\ln q\right)^2\right\}, \quad (2.3.15)$$

where $\psi^{*}\left(m_1,m_2\right) = \left[\psi^{(1)}(m_1-1)+\psi^{(1)}(m_1)\right]\left[\psi^{(1)}(m_2-1)+\psi^{(1)}(m_2)\right]$. Note that $\mathbf{V}\left\{\widehat{b}_{AD}\right\}$ is a particular case of $\mathbf{V}\left\{\widehat{b}_{MAD}\right\}$ when $m_1 = m_2$.

2.4 Monte Carlo Study

Here, the methods discussed in the previous sections are analyzed and compared for finite sample sizes. The performance of the methods is based on the empirical mean of the bias, standard deviation (sd), Mean Squared Error (MSE), size and power. In addition, the standard methods for non-cointegration tests are also considered for comparison purpose. The finite sample size investigation is divided in two parts. The first presents the empirical performance of the tests under the null hypothesis of non-cointegration for $\beta = 1.0$ in Equation 2.2.4, and the second part discusses robustness properties of the proposed tests across different values of β with a comparison to the one given in Velasco (2003).

2.4.1 Empirical results for testing non-cointegration

Let the vector $X_{1,t}, X_{2,t}$, $t = 1, ..., n$, generated by the following structure

$$X_{1,t} = X_{2,t} + \varepsilon_t$$

where the vector $(X_{2,t}, \varepsilon_t)$ was generated from

$$\begin{bmatrix} (1-B) & 0 \\ 0 & (1-B)^{1-b} \end{bmatrix} \begin{bmatrix} X_{2,t} \\ \varepsilon_t \end{bmatrix} = \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right].$$

The sample sizes considered were $n = 100, 500, 1000$ and the results are based on 3500 replications. In all cases, the 300 first observations were discarded to avoid any influence of the initial data. The parameter b assumed values $\{0, 0.1, 0.2, 0.5, 0.7, 1\}$ and the bandwidth $m = n^{0.7}$. For testing $H_0 : b = 0$ against the alternative $H_1 : b > 0$, the nominal size was fixed at 5% and $q = \{0.6, 0.8\}$ were considered for $\widehat{b}_{AD}$ and $\widehat{b}_{MAD}$, respectively. For the log determinant regression method, that is, the $(\widehat{b}_{LDR})$ estimator, r was equal to 1, which means that 3 different frequencies were included in the calculation of $\widehat{D}(\lambda_j)$ (see Equation 2.3.2).

Since the correlation between $\widehat{D}(\lambda_m)$ and $\widehat{D}(q\lambda_m)$ does not have a closed form, 2000 replications of these quantities were computed in order to obtain the estimates of the variances of $\widehat{b}_{AD}$ and $\widehat{b}_{MAD}$. From the empirical results, the sample correlations are very close for the above sample sizes across the values of $q = 0.6, 0.65, ..., 0.90, 0.95$. Some of the results are displayed in Table 2.1. Note that q was chosen not to belong to $\in (0.0, 0.5)$ to avoid loss of information from the peridogram function, that is, the estimators would have less frequencies in the calculation of the determinants of these estimators compared with $q \in (0.5, 1.0)$.

Tables 2.2 and 2.3 display results for the AD and MAD methods, respectively. Critical values for the "*t-like*"statistic, that is, the standardized statistic, were calculated from a standard Gaussian distribution. σ_n stands for the "asymptotic"standard deviations and they were calculated taking into account the mean values of correlations presented in Table 2.1.

In general, the AD and MAD procedures present quite similar results regarding the bias. Although the latter is oversized, it has smaller MSE and higher power than AD, which presents empirical size always close to 5%.

Estimated values of b are well centered around the true value in all cases considered. However, the bias increases as b becomes larger and close to one. In addition, it seems that the bias does not depend on the q value.

The power of the MAD procedure increases significantly since more frequencies are introduced in the term $\widehat{D}(\lambda_m)$. In addition, the power decreases significantly when a higher value of q is chosen. This is an expected result, since the variance depends positively on q.

As will be discussed in Section 4.2, an advantage of the AD and MAD tests over the one proposed in Velasco (2003) is that they are robust against β values.

Tabela 2.1: Simulated values for correlations used in the asymptotic variance of AD and MAD

	$\hat{b}_{AD}$				$\hat{b}_{MAD}$			
n	150	300	500	Mean	150	300	500	Mean
q								
0.60	0.4701	0.4611	0.4583	0.4631	0.6733	0.6749	0.6692	0.6725
0.65	0.4629	0.4641	0.4401	0.4557	0.6739	0.6455	0.6825	0.6673
0.70	0.4430	0.4895	0.4782	0.4702	0.6606	0.6802	0.6754	0.6721
0.75	0.4633	0.4671	0.4695	0.4666	0.6945	0.6692	0.6639	0.6759
0.80	0.4924	0.4643	0.4803	0.4790	0.6809	0.6640	0.6587	0.6678
0.85	0.4733	0.4784	0.4915	0.4811	0.6513	0.6813	0.6736	0.6687
0.90	0.4857	0.4558	0.4645	0.4686	0.6892	0.7083	0.6863	0.6946
0.95	0.4917	0.4628	0.4857	0.4801	0.6867	0.6912	0.6962	0.6914

Table 2.4 shows the critical values for the LDR test using the three sample

sizes: $n = \{100, 500, 1000\}$. These values were used to compute the results given in Table 2.5, which displays the performance of LDR method. In Table 2.5, σ_n refers to exact standard deviation calculated by .2.7.

Similar to the previous procedures, LDR method produced estimated values of b that are centered around the true value in all cases, except for $b = 1$.

The empirical size for the LDR procedure is always close to the 5%. Based on Figure 2.1, the LDR statistic completely dominates the averaged determinant statistics. Note that, the asymptotical behavior of LDR, AD and MAD, under H_1, have not been not been established yet.

In addition to the better power perfomace, there are two advantages of LDR over AD and MAD. First, the variance is known and it is easy to be calculated and the second one is that the empirical distribution for small sample sizes can be easily obtained.

Figure 2.2 plots the "*t-like*"densities for all tests, AD, MAD and LDR and the standard Gaussian density. For AD and MAD, the densities were computed only for $q = 0.6$, since this presented better power than $q = 0.8$.

One can see that the empirical densities of AD are closer to the Normal density curve than the MAD. As expected, the density of LDR statistic is very close to the standard Gaussian density, even for $n = 100$.

Based on the previous discussions of the advantages of the LDR test over the AD and MAD methods, the first method is now compared with the residual based tests given in Dittmann (2000). The tests are: GPH, Lagrange Multiplier (LM)[4], Modified Rescaled Range (MRR)[5], Phillips-Perron ρ-test (PP$_\rho$), Phillips-Perron

[4]See Lobato and Robinson (1998) for details
[5]See Lo (1991) for details

t-test (PP_t) and Augmented Dickey-Fuller (ADF) [6].

In order to make the above comparison, the same experiments conducted by Dittmann (2000) were also used here to obtain the empirical size and power of the LDR test and these are displayed in Table 2.6, which also includes the results from Table A3 in Dittmann (2000).

In general, it can be observed that LDR plays an intermediate role when compared to other tests. When sizes are evaluated, that is, when the simulated series are ARIMA(1,1,0) and ARIMA(0,1,1) models, LDR completely dominates the frequency domain GPH and LM tests by presenting less oversized significance levels. In the case of the time domain tests, LDR displayed better sizes than PP_ρ and PP_t in most of the cases. The ADF and MRR tests show the best size performance among the evaluated tests. However, they have less power when $n \geq 250$ and when d is close to the null hypothesis of non-cointegration. The remaining tests possess higher power in the majority of cases. The only exceptions are the PP_ρ and PP_t tests that lose power when models are near to the null hypothesis for sample sizes equal or greater than 250.

[6]See Hamilton (1994) for details of Phillips-Perron ρ-test, Phillips-Perron t-test and Augmented Dickey-Fuller test

Tabela 2.2: Estimates, size and power for the AD method at 5% significance level

q = 0.6		b=0	b=0.1	b=0.2	b=0.5	b=0.7	b=1
n = 100	Mean	0.0177	0.1052	0.2086	0.4627	0.6373	0.9082
σ_n = 0.4116	Sd	0.4186	0.4228	0.4292	0.4732	0.4902	0.5558
	MSE	0.1755	0.1788	0.1843	0.2253	0.2441	0.3173
	Rejection	0.0540	0.0900	0.1374	0.3160	0.4611	0.6494
n = 500	Mean	0.0066	0.1073	0.1976	0.4918	0.6804	0.9581
σ_n = 0.2312	Sd	0.2384	0.2452	0.2485	0.2762	0.3034	0.3232
	MSE	0.0569	0.0602	0.0617	0.0763	0.0924	0.1062
	Rejection	0.0597	0.1351	0.2271	0.6474	0.8383	0.9646
n = 1000	Mean	0.0044	0.1056	0.2038	0.4987	0.6880	0.9661
σ_n = 0.181	Sd	0.1839	0.1875	0.1921	0.2162	0.2334	0.2504
	MSE	0.0338	0.0352	0.0369	0.0467	0.0546	0.0638
	Rejection	0.0520	0.1534	0.3111	0.8246	0.9583	0.9974
q = 0.8		b=0	b=0.1	b=0.2	b=0.5	b=0.7	b=1
n = 100	Mean	0.0446	0.1013	0.1999	0.4756	0.6551	0.8945
σ_n = 0.9512	Sd	0.9862	0.9750	1.0047	1.0977	1.1613	1.2288
	MSE	0.9743	0.9504	1.0092	1.2052	1.3503	1.5206
	Rejection	0.0626	0.0654	0.0857	0.1557	0.2111	0.2874
n = 500	Mean	-0.0052	0.1043	0.2133	0.4949	0.6798	0.9738
σ_n = 0.5343	Sd	0.5389	0.5503	0.5690	0.6067	0.6526	0.7101
	MSE	0.2903	0.3027	0.3239	0.3680	0.4262	0.5048
	Rejection	0.0480	0.0791	0.1251	0.2637	0.3826	0.5491
n = 1000	Mean	0.0113	0.0976	0.2103	0.4940	0.6797	0.9790
σ_n = 0.4182	Sd	0.4255	0.4243	0.4392	0.4842	0.5057	0.5338
	MSE	0.1811	0.1799	0.1929	0.2344	0.2560	0.2853
	Rejection	0.0543	0.0860	0.1377	0.3500	0.4923	0.7146

Tabela 2.3: Estimates, size and power for the MAD method at 5% significance level

q = 0.6		b=0	b=0.1	b=0.2	b=0.5	b=0.7	b=1
n = 100	Mean	0.0118	0.1133	0.2118	0.4732	0.6548	0.8929
σ_n = 0.2972	Sd	0.3088	0.3222	0.3328	0.3747	0.4035	0.4386
	MSE	0.0955	0.1040	0.1109	0.1411	0.1648	0.2037
	Rejection	0.0683	0.1277	0.1966	0.4786	0.6417	0.8200
n = 500	Mean	0.0093	0.1033	0.2030	0.4946	0.6758	0.9538
σ_n = 0.1669	Sd	0.1763	0.1795	0.1925	0.2148	0.2308	0.2619
	MSE	0.0312	0.0322	0.0371	0.0462	0.0539	0.0707
	Rejection	0.0723	0.1720	0.3463	0.8497	0.9591	0.9966
n = 1000	Mean	0.0033	0.1051	0.1988	0.4961	0.6875	0.9690
σ_n = 0.1307	Sd	0.1345	0.1419	0.1476	0.1726	0.1880	0.2042
	MSE	0.0181	0.0201	0.0218	0.0298	0.0355	0.0426
	Rejection	0.0609	0.2134	0.4540	0.9563	0.9954	1.0000

q = 0.8		b=0	b=0.1	b=0.2	b=0.5	b=0.7	b=1
n = 100	Mean	0.0339	0.1329	0.2321	0.4738	0.6482	0.8963
σ_n = 0.6803	Sd	0.6901	0.7172	0.7386	0.7927	0.8413	0.9111
	MSE	0.4772	0.5153	0.5464	0.6289	0.7103	0.8407
	Rejection	0.0671	0.0954	0.1174	0.1983	0.2766	0.3883
n = 500	Mean	0.0149	0.1236	0.2118	0.4859	0.6843	0.9584
σ_n = 0.3821	Sd	0.3911	0.4036	0.4214	0.4593	0.4923	0.5432
	MSE	0.1532	0.1634	0.1777	0.2111	0.2425	0.2967
	Rejection	0.0643	0.1063	0.1637	0.3589	0.5297	0.7160
n = 1000	Mean	0.0042	0.1042	0.1979	0.4952	0.6873	0.9595
σ_n = 0.2991	Sd	0.3093	0.3117	0.3261	0.3716	0.3876	0.4285
	MSE	0.0957	0.0972	0.1063	0.1381	0.1504	0.1852
	Rejection	0.0657	0.1117	0.1846	0.4911	0.6857	0.8631

Tabela 2.4: Critical values for LDR $t - like$ statistic

Significance Level	90%	95%	99%
n = 100	1.2804	1.6792	2.4980
n = 500	1.2845	1.6801	2.4494
n = 1000	1.2890	1.6796	2.4380

Tabela 2.5: Estimates, size and power for the LDR method at 5% significance level

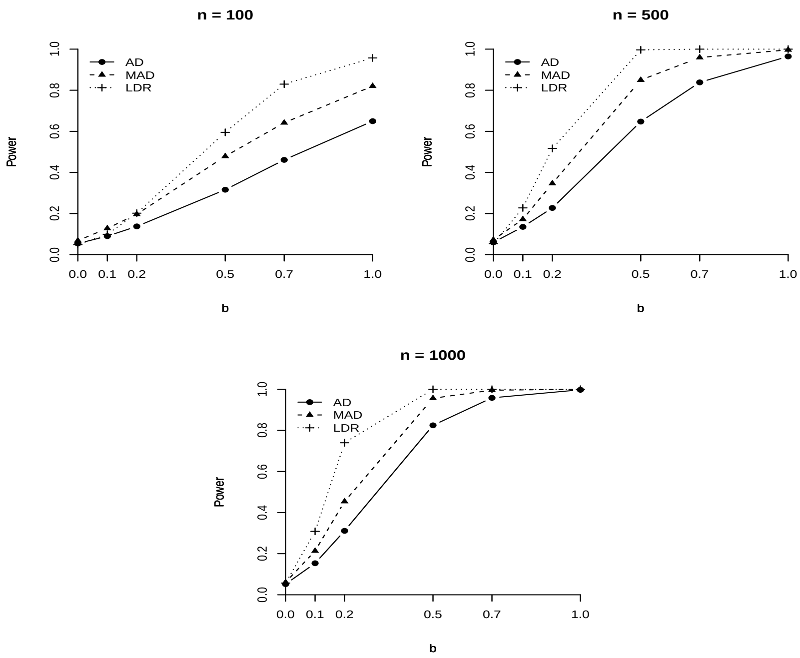
		b=0	b=0.1	b=0.2	b=0.5	b=0.7	b=1
n=100	mean	0.0018	0.1020	0.2104	0.4764	0.6487	0.8689
$\sigma_n = 0.2433$	sd	0.2424	0.2401	0.2467	0.2505	0.2536	0.2736
	mse	0.0588	0.0576	0.0610	0.0633	0.0669	0.0920
	Rejection	0.0483	0.0991	0.2017	0.5951	0.8297	0.9571
n=500	mean	-0.0008	0.1056	0.1978	0.4830	0.6562	0.8545
$\sigma_n = 0.1134$	sd	0.1134	0.1129	0.1141	0.1145	0.1205	0.1583
	mse	0.0129	0.0128	0.0130	0.0134	0.0164	0.0462
	Rejection	0.0531	0.2274	0.5169	0.9963	1.0000	1.0000
n=1000	mean	0.0023	0.1013	0.1994	0.4854	0.6639	0.8560
$\sigma_n = 0.0859$	sd	0.0879	0.0871	0.0847	0.0871	0.0970	0.1385
	mse	0.0077	0.0076	0.0072	0.0078	0.0107	0.0399
	Rejection	0.0571	0.3086	0.7397	1.0000	1.0000	1.0000

Tabela 2.6: Size and power comparison at 5% significance level

		Power					Size									
		I(d) processes with $d =$					ARIMA(1,1,0) with $\phi =$					ARIMA(0,1,1) with $\theta =$				
Sample Size	Test	0.1	0.3	0.5	0.7	0.9	-0.474	-0.412	-0.333	-0.231	-0.091	0.718	0.525	0.382	0.245	0.092
100	GPH	98.93	97.05	84.02	45.15	11.11	10.15	10.53	10.12	8.56	6.85	86.37	46.98	24.98	13.18	6.71
	LM	99.89	99.25	92.96	60.92	13.35	16.8	17.48	16.06	13.65	8.13	96.53	71.84	44.08	22.43	8.68
	MRR	15.32	26.15	30.82	24.19	9.58	2.2	3.16	4.27	5.33	5.87	5.34	5.98	6.08	6.6	5.78
	PP_ρ	100	100	96.26	52.57	11.22	23.85	20.06	14.7	10	6.54	89.32	51.25	25.96	13.84	6.86
	PP_t	100	100	95.28	48.83	10.45	23.82	20.01	14.78	9.79	6.42	89.32	50.77	25.39	13.17	6.68
	ADF	98.33	93.44	70.88	36.33	9.46	4.73	5.82	6.72	6.83	5.73	41.85	17.77	11.06	9.06	5.74
	LDR	93.59	82.47	60.53	30.44	10.38	9.82	8.98	8.59	7.64	6.13	72.94	36.53	19.88	10.96	6.46
250	GPH	99.99	99.99	99.78	85.07	18.63	10.43	10.55	10.88	9.88	7.57	96.77	61.09	31.11	16.39	8.01
	LM	100	100	99.96	94.4	25.25	14.83	15.14	15.07	13.88	8.97	99.26	82.4	51.22	26.86	10.02
	MRR	61.49	72.15	68.09	46.55	13.91	2.73	3.32	4.51	5.21	6.1	8.83	5.46	5.85	6.04	5.75
	PP_ρ	100	100	99.98	78.65	15.36	20.66	16.18	11.95	9.49	6.58	89.39	47.19	22.2	11.4	6.88
	PP_t	100	100	99.94	75.98	13.84	20.66	15.79	11.72	9.16	6.18	89.22	46.66	21.55	10.87	6.52
	ADF	99.97	99.48	89.75	50.78	12.3	4.59	4.54	5.2	5.76	5.89	23.96	9.89	7.32	6	5.93
	LDR	99.9	99.36	93.28	58.81	15.06	8.67	8.16	7.93	7.1	5.84	88.43	43.18	20.75	10.8	6.46
500	GPH	100	100	100	98.38	30.49	10.18	9.69	9.81	9.41	7.62	93.98	68.8	34.95	17.64	8.55
	LM	100	100	100	99.77	41.43	13.6	13.04	13.8	12.98	10.2	99.62	86.5	55.26	28.7	11.21
	MRR	95.61	97.82	94.46	68.75	17.77	3.41	4.22	4.56	5.27	5.8	13.58	6.5	5.75	5.98	5.58
	PP_ρ	100	100	100	90.16	19.13	17.95	13.96	10.21	7.73	5.75	87.16	41.21	18.9	9.8	6.47
	PP_t	100	100	100	88.3	16.82	17.17	13.23	9.91	7.14	5.41	86.81	40.59	18.15	9.32	6.22
	ADF	100	99.97	97.81	65.72	14.45	4.44	4.75	5.02	4.64	5.03	16.34	8.09	6.38	5.94	5.6
	LDR	100	99.99	99.6	82.56	21.1	7.56	7.64	7.34	6.91	5.67	92.81	42.08	18.71	10.32	6
1000	GPH	100	100	100	99.99	50.39	9.79	9.65	9.98	9.72	7.98	98.91	75.47	38.43	19.31	8.9
	LM	100	100	100	100	66.16	13	12.78	13.2	12.93	11.12	99.66	88.94	59.45	60.51	12.79
	MRR	99.91	99.98	99.79	88.64	23.93	4.11	4.89	5.19	5.7	5.74	16.92	6.82	6.07	6.11	5.88
	PP_ρ	100	100	100	96.29	24.24	15.12	11.23	8.97	7.32	6.03	84.51	35.51	17.93	8.47	6
	PP_t	100	100	100	95.46	21.54	14.55	10.66	8.55	7	6.02	84	34.66	14.21	8.29	5.96
	ADF	100	100	99.76	79.21	17.95	5.05	4.79	4.56	4.85	5.43	12.88	6.85	5.82	5.67	5.4
	LDR	100	100	100	96.4	29.95	7.29	7.03	6.01	6.43	5.9	94.71	39.67	16.69	9.38	5.98

Note: Values from GPH, LM, MRR, PP_ρ, PP_t and ADF are from Table A3, p. 638 by Dittmann (2000).

Figura 2.1: Power of AD, MAD and LDR at 5% significance level

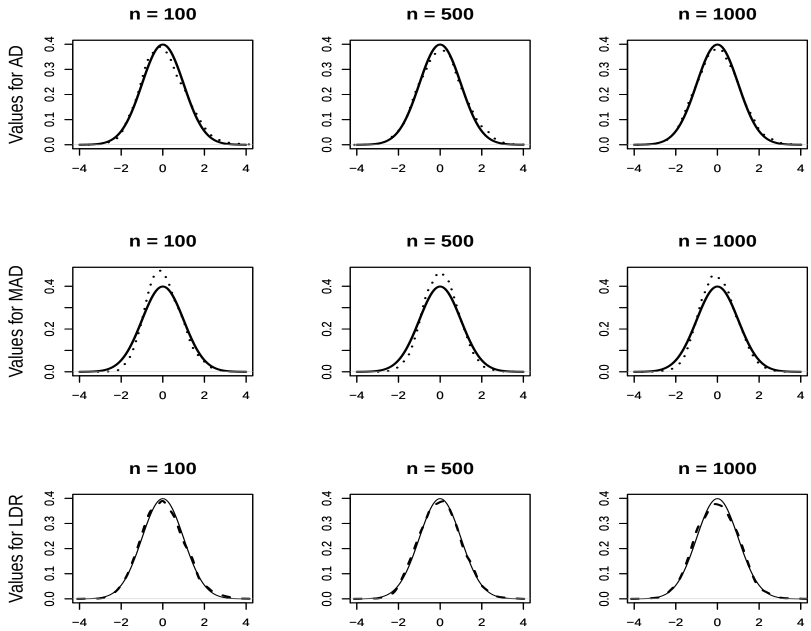


2.4.2 Robustness to different values of cointegration declivity (β) - A comparison with log coherence regression

This subsection discusses the robustness properties of LDR, AD, MAD and the estimator given in Velasco (2003) over the declivity (β) values. Velasco (2003) proposed a method to estimate parameter b and test the null of cointegration based on a regression of logged squared coherence between the pair of series $(X_{1,t}, X_{2,t})$ on logged Fourier frequencies. This procedure exploits the fact that near zero frequency the squared coherence for cointegrated series converges to 1.

The LCR method is built for a non differentiated vector $(X_{1,t}, X_{2,t})$ where $X_{i,t} \sim I(d)$, $i \in \{1, 2\}$, with $d \in (0, 1.5)$. In the context that the series are non-stationary, the author suggests to use data tapering for controlling the peri-

Figura 2.2: Empirical densities of standardized statistics for AD, MAD and LDR



Note: Solid lines are Gaussian density functions and dashed lines are the empirical densities.

odogram bias. He shows the consistency of his estimator when $0 < b \leq d$, that is, when the vector are cointegrated. When the vector $(X_{1,t}, X_{2,t})$ is non-stationary its spectral density matrix, denoted by $\mathbf{H}(\lambda)$, becomes , as mentioned early, a pseudo-spectral density matrix. $\mathbf{H}(\lambda)$ is given by:

$$\mathbf{H}(\lambda) = \lambda^{-2d} \begin{pmatrix} G_{xx} & G_{xe}\lambda^{b} \\ G_{ex}\lambda^{b} & G_{ee}\lambda^{2b} \end{pmatrix} (1 + O(\lambda^2)) \; as \; \lambda \to 0^+ \qquad (2.4.1)$$

for some constants $|G_{ab}| < \infty, a, b \in \{x, e\}$ where the matrix $G = \{G_{ab}\}$ is Hermitian and nonsingular. The pseudo-spectral density for $X_{1,t}$ and the cross pseudo spectral density for $X_{1,t}$ and $X_{2,t}$ can be written, respectively as:

$$f_{X_1X_1}(\lambda) = \beta^2 f_{X_2X_2}(\lambda) + f_{\varepsilon\varepsilon}(\lambda) + 2\beta \mathrm{Re} f_{\varepsilon X_2}(\lambda) \sim \beta^2 G_{xx}\lambda^{-2d} \; as \; \lambda \to 0^+ \qquad (2.4.2)$$

and

$$f_{X_2X_1}(\lambda) = \beta f_{X_2X_2}(\lambda) + f_{X_2\varepsilon}(\lambda) \sim \beta G_{xx}\lambda^{-2d} \; as \; \lambda \to 0^+ \tag{2.4.3}$$

where $\mathrm{Re}(z)$ means the real part of a complex number z. The squared coherence $|\kappa_{X_1X_2}(\lambda)|^2$ between $X_{1,t}$ and $X_{2,t}$ can be written as:

$$|\kappa_{X_1X_2}(\lambda)|^2 = 1 - \frac{f_{\varepsilon\varepsilon}(\lambda)}{f_{X_1X_1}(\lambda)} + \frac{|f_{\varepsilon X_2}(\lambda)|^2}{f_{X_1X_1}(\lambda)\, f_{X_2X_2}(\lambda)}. \tag{2.4.4}$$

Replacing the approximations in Equations 2.4.1, 2.4.2 and 2.4.3 into the above expression and taking logs yield:

$$\ln\left(1 - |\kappa_{X_1X_2}(\lambda)|^2\right) \sim \ln G_H + 2b\ln\lambda \; as \; \lambda \to 0^+ \tag{2.4.5}$$

where $0 < G_H < \infty$ and $G_H = \frac{G_{ee}}{\beta^2 G_{xx}}\left[1 - \frac{|G_{ex}|^2}{G_{xx}G_{ee}}\right]$.

Using consistent estimates of $|\kappa_{X_1X_2}(\lambda)|^2$, the parameter b can be estimated by regressing $\ln\left(1 - |\hat{\kappa}_{X_1X_2}(\lambda)|^2\right)$ on logged values of frequencies λ around the origin. This estimator is denoted here by the log coherence regression (LCR).

Despite LCR differs from AD, MAD and LDR in terms of the null hypothesis, the method also estimates the same parameter, b, which makes the comparison meaningful.

Although LCR asymptotically does not depend on β, this parameter can be influential on small samples. Simulation results presented in this section show that, in fact, the values of β can affect the estimated values for b. On the other hand, the determinant based estimators, LDR, AD and MAD, are unaffected by the choice of β.

To evaluate the robustness properties of the estimators over β values, a Monte

Carlo experiment was conducted with $CI(1,1)$. Parameter β varies in the range $\{1, 0.5, 0.125, 0.065\}$. The system described in Equation 2.2.1 was simulated with innovations following a Gaussian white noise process with $\sigma_\eta^2 = 1$, $\sigma_{w_1}^2 = 1$, $\sigma_{w_2}^2 = 0$ and $\beta_2 = 1$ (which makes $X_{2,t} = T_t$ and $cov(\varepsilon_t, X_{2,t}) = 0$). In all replications the sample size was fixed at $n = 256$ in order to keep comparability to the models simulated by Velasco (2003)[7].

Bandwidth for LCR was fixed at $m = 36$ since this value has displayed the best results, in terms of mean squared error, for most models considered in Velasco (2003). For AD, MAD and LDR, the bandwidth was fixed at $m = n^{0.7}$, which was the usual choice of the present work, while the parameter q was set equal to 0.6 for AD and MAD. Table 2.7 presents the mean of the estimates, the standard deviation and the mean squared error.

As displayed in Table 2.7, the LCR is not robust to different values of the cointegration vector β. The bias and mean squared errors of estimates increase with decreasing β. The AD and MAD seem to be the less biased estimators and insensitive to variations in β while LDR, although biased, produced the best performance in terms of mean squared error and is also insensitive to the changes in the cointegration vector. These simulations illustrate that LCR is not a very reliable estimator for small samples.

[7]In the LCR models simulated here a Zhurbenko data taper of order 2 was used as suggested by the author. It means that the fast Fourier transform used to calculate the periodogram assumes the form: $\left|\sum_{t=1}^{n} h_t X_t e^{-i\lambda t}\right|^2$ were h_t is the data taper such that $h_t = 1 - \left|\frac{t-n/2}{n/2}\right|$. Also, estimates of 2.4.4 are using previously defined spectral estimates with uniform weights over 3 Fourier frequencies. See Velasco (2003) for details

Tabela 2.7: Robustness properties of the methods to β variations

β	Statistic	$\hat{b}_{LCR}$	$\hat{b}_{AD}$	$\hat{b}_{MAD}$	$\hat{b}_{LDR}$
1	Mean	0.8399	0.9482	0.9719	0.8505
	Sd	0.2195	0.4147	0.3253	0.1882
	MSE	0.0738	0.1746	0.1066	0.0578
0.5	Mean	0.6718	0.9353	0.9655	0.8531
	Sd	0.1991	0.4063	0.3287	0.1938
	MSE	0.1474	0.1692	0.1092	0.0591
0.125	Mean	0.2190	0.9199	0.9722	0.8570
	Sd	0.1550	0.4052	0.3286	0.1867
	MSE	0.6341	0.1706	0.1087	0.0553
0.065	Mean	0.0814	0.9502	0.9705	0.8529
	Sd	0.1256	0.4125	0.3298	0.1844
	MSE	0.8596	0.1726	0.1096	0.0556

2.5 Application

The methodologies previously described were applied to monthly observations of logged stock values for United States markets (Dow Jones Industrial Average Index - DJ) and United Kingdom markets (Financial Times Stock Exchange 100 - FTSE). Data ranges from January 1985 to May 2014. Looking at Figure 2.3 it can be seen that stock values in US and UK markets seem to share some long-run relationship.

Firstly, GPH was employed in both series to estimate d using two different values of bandwidth $n^{0.5}$ and $n^{0.7}$. For the DJ series the estimated values of d were 0.9281 (sd = 0.1942) and 0.9878 (sd = 0.0941) for $n^{0.5}$ and $n^{0.7}$, respectively, while, for the FTSE series, estimated values of d were 0.9490 (sd = 0.1941) and 1.0181 (sd = 0.0941) for $n^{0.5}$ and $n^{0.7}$, respectively. These results were used to calculate the *t-like* statistic for the GPH unit root test (see Santander et al. (2003)) in which the null hypothesis is H_0 : *the series is I(1)* versus H_1 : *the series is I(d)*, $d < 1$. In addtion, the ADF and PP tests were also implemented. From Table 8 it can be seen that these tests suggest that both series have a unit root.

The next step was to test the null of non-cointegration using LDR, AD, MAD and also using GPH and EG[8]. The last two tests were performed on the residuals of the regression equation for DJ and FTSE. Values for AD and MAD were obtained using $q = 0.6$. Estimated values for parameter b using LDR, AD, MAD and GPH (which was obtained using $1 - \hat{d}_{res}$, where $\hat{d}_{res}$ is the estimated d for regression residuals) can be seen in Table 2.9. Note that the EG test does not estimate the parameter b.

The results in Table 2.9 indicated that the series are not cointegrated, which is in accordance with the most of the empirical evidence discussed in the literature, that is, most of the international stock prices analysed are not pairwise cointegrated (see Aloy et al. (2013) or Kanas (1998)).

2.6 Conclusion

The present work investigates the properties of the determinant of the spectral density matrix close to the origin for bivariate cointegrated series and proposes methods to test the null hypothesis of non-cointegration based on these properties.

The determinant of the spectral density matrix for the first difference series is a power function of the parameter b, which determines the reduction in the order of integration of the error series. Two different statistics were considered: the log determinant regression and semiparametric averaged determinant estimator.

Monte Carlo simulations showed that the methods presented here possess, in

[8]Critical values for GPH test can be found in Santander et al. (2003). For EG test see MacKinnon (1991) while for AD and MAD asymptotic standard Gaussian values were used

Tabela 2.8: Values of unit root test statistic and critical values

	GPH[a]($n^{0.5}$)	GPH[a]($n^{0.7}$)	ADF	PP
DJ	-0.3702	-0.1297	2.2928	-1.8714
FTSE	-0.2628	0.1923	1.5746	-2.0234
Critical Values ($\alpha = 5\%$)	-1.61	-1.58	-1.95	-2.87
Critical Values ($\alpha = 10\%$)	-1.16	-1.18	-1.62	-2.57

[a] Critical values for GPH test can be found in Santander et al. (2003)

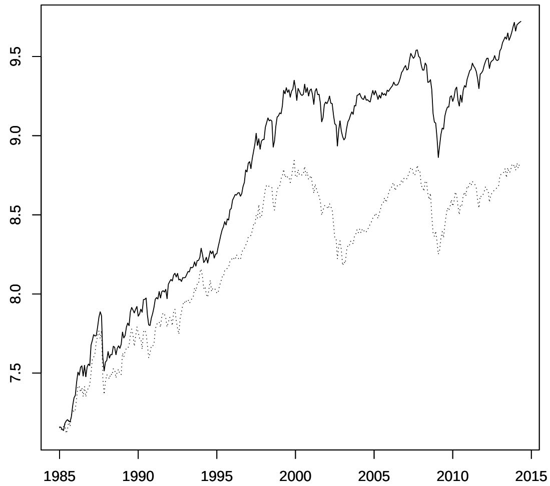
Tabela 2.9: Estimates for b and test statistic for non-cointegration between DJ and FTSE

	LDR		AD		MAD		GPH[ab]		EG
Bandwidth	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	-
$\hat{b}$	-0.14	0.02	-0.45	0.30	-0.24	0.14	-0.14	0.03	-
Standard Deviations	0.29	0.13	0.51	0.27	0.36	0.19	0.19	0.09	-
t-like Statistic	-0.49	0.18	-0.88	1.11	-0.67	0.74	-0.74	0.32	-2.27
Critical Values ($\alpha = 5\%$)	1.68	1.68	1.64		1.64		2.24	2.11	-2.87
Critical Values ($\alpha = 10\%$)	1.29	1.29	1.28		1.28		1.78	1.67	-2.57

[a] In order to keep comparability with LDR, AD and MAD, critical values were adjusted to test $H_0 : 1 - d_{res} = 0$ versus $H_1 : 1 - d_{res} > 0$

[b] Critical values for GPH test can be found in Santander et al. (2003)

Figura 2.3: Logarithm of stock values, Dow Jones (solid) and FTSE 100 (dotted)



general, correct size. The log determinant regression showed good power for moderate sample size. In addition, Monte Carlo simulations also showed that both methods are insensitive to the choice of cointegration relationship (the β parameter). Some properties of the above statistics under the null of non-cointegration were discussed and the current research is investigating the asymptotic properties under the alternative hypothesis as well as the extension of those proposals to a higher dimension vector allowing more than one cointegration relationship. In addition, a version of these tests which is robust to the presence of outliers is being investigated.

Acknowledgements

V.A. Reisen was partially supported by CNPq-Brazil, FAPES Foundation and CAPES.

G.C. Franco was partially supported by CNPq-Brazil (grant number: 305646/2012-5), and also by FAPEMIG Foundation (grant number: PPM-00021-14).

.1 Technical Lemmas

Before proving the propositions, one should consider the following Lemmas:

Lemma 1. *(Hurvich and Beltrao (1994) pg. 301)*
Let $Z_j = \ln(2 - 2\cos\lambda_j)$ *and* $\tilde{Z}_j = Z_j - \bar{Z}$, $j = 1, ..., m$, *where* $\bar{Z}$ *is the mean of* Z_j. *Thus,* $\sum_{j=l}^{m} \tilde{Z}_j^{\,2} = m + o(m)$.

Lemma 2. *Let* $Z_j = \ln(2 - 2\cos\lambda_j)$ *and* $\tilde{Z}_j = Z_j - \bar{Z}$, $j = 1, ..., m$, *where* $\bar{Z}$ *is the mean of* Z_j. *Thus,* $Z_j = O(\ln m)$.

Demonstração. Hurvich and Beltrao (1994) state that $Z_j = \ln j - \frac{1}{m}\ln m! + o(1)$, where the first term, $\ln j = O(\ln m)$ $\forall j = 1, ..., m$. The following limit,

$$\lim_{m\to\infty} \frac{1}{m} \ln m!$$

is an indeterminate form of the type $\frac{\infty}{\infty}$. By the L'Hospital's rule, the previous limit can be rewritten as: $\lim_{m\to\infty} \psi^{(0)}(m+1)$. Equation 6.3.18 from Abramowitz and Stegun (1972), pg. 259, states that: $\psi^{(0)}(z) \sim \ln z - \frac{1}{2z} - \frac{1}{12z^2} + \frac{1}{120z^4} - \frac{1}{256z^6} + ...$ as $z \to \infty$. As a result, $\psi^{(0)}(z) = O(\ln z)$ and, consequently, $Z_j = O(\ln m)$. □

Lemma 3. *Let* $\psi^{(1)}(u)$ *be the Polygamma function of order 1, that is,* $\psi^{(1)}(u) = \frac{d^2 \ln\Gamma(u)}{du^2}$. *Then, for* $z \in \mathbb{N}^*$, $\lim_{z\to\infty} \psi^{(1)}(z) = 0$.

Demonstração. Jeffrey and Zwillinger (2007), pg.905, state that for any $z \in \mathbb{N}^*$, $\psi^{(1)}(z) = \frac{\pi^2}{6} - \sum_{k=1}^{z-1} \frac{1}{k^2}$. Thus, $\lim_{z\to\infty} \psi^{(1)}(z) = \frac{\pi^2}{6} - \lim_{z\to\infty} \sum_{k=1}^{z-1} \frac{1}{k^2}$. In addition, Jeffrey and Zwillinger (2007), pg.8, state that $\lim_{z\to\infty} \sum_{k=1}^{z-1} \frac{1}{k^2} = \frac{\pi^2}{6}$, as a result, $\lim_{z\to\infty} \psi^{(1)}(z) = 0$.

□

.2 Proof of Proposition 1

1. Under Assumption 1, the determinant $D(\lambda_j)$ is constant across different frequencies, that is, $D(\lambda_j) = \Lambda$, for all $\lambda_j \in (0, 2\pi]$, where Λ is a positive constant. In this case, for a fixed r, Goodman (1963) points out that the distribution of the random variable $\mathscr{D}_j = \ln\left[4\,(2r+1)^2\,\widehat{D}_r(\lambda_j)/\Lambda\right]$ is the same of $ln(\chi^2_{(4r+2)}\chi^2_{(4r)})$ where $\chi^2_{(4r+2)}$ and $\chi^2_{(4r)}$ are chi-squared random variables with $(4r+2)$ and $4r$ degrees of freedom, respectively. In addition, the characteristic function of $\mathscr{D}_j$ is:

$$\phi_{\mathscr{D}}(t) = 4^{it}\frac{\Gamma(2r+1+it)\Gamma(2r+it)}{\Gamma(2r+1)\Gamma(2r)}. \tag{.2.1}$$

Taking the first derivative of $\phi_{\mathscr{D}}(t)$ evaluated at $t = 0$ yields:

$$\frac{d\phi_{\mathscr{D}}(t)}{dt}\bigg|_{t=0} = (\psi^{(0)}(2r+1) + \psi^{(0)}(2r) + \ln 4)i. \tag{.2.2}$$

Then, $\mathbf{E}\left[\mathscr{D}_j\right] = \psi^{(0)}(2r+1) + \psi^{(0)}(2r) + \ln 4$. Since $\mathscr{D}_j = \ln\left[\frac{\widehat{D}_r(\lambda_j)}{\Lambda}\right] + \ln\left[4\,(2r+1)^2\right]$, one can see that:

$$\mathbf{E}\left[\ln\left[\frac{\widehat{D}_r(\lambda_j)}{\Lambda}\right]\right] = \psi^{(0)}(2r+1) + \psi^{(0)}(2r) - 2\ln(2r+1). \tag{.2.3}$$

Let $\zeta_j = \left\{\ln\left[\frac{\widehat{D}_r(\lambda_j)}{\Lambda}\right] - c(r)\right\}$, where $c(r) = \mathbf{E}\left[\ln\left[\frac{\widehat{D}_r(\lambda_j)}{\Lambda}\right]\right]$. It is easy to see that ζ_j is a random variable such that $\mathbf{E}\left[\zeta_j\right] = 0$. In addition, if Assumption 1 holds, then term $\ln\frac{G(\lambda_j)}{G(0)}$ in Equation 2.3.5 is equal to zero since $G(\lambda_j) = G(0)$, $\forall \lambda_j$. Let $a = \ln G(0) + c(r)$. It can be seen that:

$$\mathbf{E}\left[\hat{b}_{LDR}\right] = \mathbf{E}\left[\left(\sum_{j=l}^{m} \tilde{Z}_j^{\,2}\right)^{-1} \sum_{j=l}^{m} \tilde{Z}_j(a + bZ_j + \zeta_j)\right] =$$

$$b + \left(\sum_{j=l}^{m} \tilde{Z}_j^{\,2}\right)^{-1} \sum_{j=l}^{m} \tilde{Z}_j \mathbf{E}\left[\zeta_j\right] = b. \quad (.2.4)$$

Moreover, under Assumption 1, $b = 0$. Thus, $\mathbf{E}\left[\hat{b}_{LDR}\right] = 0$.

2. The variance $\mathbf{V}\left[\hat{b}_{LDR}\right]$ is equal to the variance of the term $\left(\sum_{j=l}^{m} \tilde{Z}_j^{\,2}\right)^{-1} \sum_{j=l}^{m} \tilde{Z}_j \zeta_j$. Under Assumption 1, the set of random variables ζ_j, $j = 1, ..., m$ is independent and identically distributed. Therefore, $\mathbf{V}\left[\hat{b}_{LDR}\right]$ is:

$$\mathbf{V}\left[\hat{b}_{LDR}\right] = \frac{\mathbf{V}\left[\zeta\right]}{\sum_{j=l}^{m} \tilde{Z}_j^{\,2}}, \quad (.2.5)$$

where $\mathbf{V}\left[\zeta\right]$ is the variance of ζ_j for all j which is the same variance of the random variable $\mathscr{D}_j$. In turn, the variance of $\mathscr{D}_j$ can be obtained from the first and second derivatives of function $\phi_{\mathscr{D}}(t)$. The second derivative of function in Equation .2.1 at $t = 0$ is:

$$\left.\frac{d^2\phi_{\mathscr{D}}(t)}{dt^2}\right|_{t=0} = -\left[2\ln 4 + \psi^{(0)}(2r)^2 + 2\psi^{(0)}(2r+1)\left(\ln 4 + \psi^{(0)}(2r+1)\right)\right.$$
$$\left. + \psi^{(0)}(2r+1)\left(\ln 16 + \psi^{(0)}(2r+1)\right) + \psi^{(1)}(2r) + \psi^{(1)}(2r+1)\right]. \quad (.2.6)$$

From Equations .2.2 and .2.6, the variance of $\mathscr{D}$ is $\psi^{(1)}(2r+1) + \psi^{(1)}(2r)$.

Replacing the quantity $\psi^{(1)}(2r+1)+\psi^{(1)}(2r)$ in Equation .2.5 yields:

$$\mathbf{V}\left[\hat{b}_{LDR}\right] = \frac{\psi^{(1)}(2r+1)\psi^{(1)}(2r)}{\sum\limits_{j=l}^{m} \tilde{Z}_j^{\,2}}. \tag{.2.7}$$

3. The proof of this part follows immediately from Lemma 1.

□

.3 Proof of Proposition 2

Let the sequence of random variables U_{mj}, $j \leq m$ form the following triangular array:

$$\begin{aligned} &U_{1,1};\\ &U_{2,1}, U_{2,2};\\ &\cdots\cdots\cdots\cdots\cdots;\\ &U_{m,1}, U_{m,2}\cdots, U_{m,m}, \end{aligned} \tag{.3.1}$$

where the random variables in each row of .3.1 are independent and for each m and j,

$$\mathbf{E}\left\{U_{mj}\right\} = 0 \tag{.3.2}$$

and $\mathbf{V}\left\{U_{mj}\right\} = \sigma^2_{mj}$ such that $\sigma^2_{mj} < \infty$ and

$$s^2_m = \sum_{j=1}^{m} \sigma^2_{mj} = 1. \tag{.3.3}$$

Theorem 7.2.1 given in Chung (2001) states that, for a triangular array where the conditions established by Equations .3.2 and .3.3 hold, the sum $S_m = \sum_{j=1}^{m} U_{mj} \xrightarrow{d} N(0,1)$, as $m \to \infty$.

In order to prove Proposition 2, one can observe that term $\left(\sum_{j=l}^{m} \tilde{Z}_j^{\,2}\right)^{-1} \sum_{j=l}^{m} \tilde{Z}_j \zeta_j$, under Assumptions 1 and 2, satisfies Equations .3.2 and .3.3. For all $j \leq m$, let

$$U_{1j} = U_{2j} = \cdots = U_{mj} = \frac{\tilde{Z}_j \zeta_j}{\sqrt{\sum_{j=l}^{m} \tilde{Z}_j^{\,2} \left(\psi^{(1)}(2r+1) + \psi^{(1)}(2r)\right)}}. \tag{.3.4}$$

Since the random variables U_{mj} are the same across rows, the subscript m can be dropped. By the definition of U_j one can see directly that $\mathbf{E}\{S_m\} = 0$. By Lemmas 1 and 2,$\mathbf{V}\{U_j\} = \frac{\tilde{Z}_j^{\,2}}{\sum_{k=l}^{m} \tilde{Z}_k^{\,2}} < \infty$. Therefore, $\mathbf{V}\{S_m\} = \mathbf{V}\left\{\sum_{j=1}^{m} U_j\right\} = \sum_{j=l}^{m} \frac{\tilde{Z}_j^{\,2}}{\sum_{k=l}^{m} \tilde{Z}_k^{\,2}} = 1$. As a result, $\sqrt{\frac{\sum_{j=l}^{m} \tilde{Z}_j^{\,2}}{\psi^{(1)}(2r+1)+\psi^{(1)}(2r)}}\hat{b}_{LDR} \xrightarrow{d} N(0,1)$ as $m \to \infty$.

□

.4 Proof of Proposition 3

1. Under Assumption 1, $D(\lambda) = \Lambda$, for all $\lambda \in [0, 2\pi)$, where Λ is a positive constant. Thus, $D(\lambda_j) = D(q\lambda_k) = \Lambda$ for all $j, k = 1, ..., m$. Therefore:

$$\mathbf{E}\left[\hat{b}_{AD}\right] = \left\{\mathbf{E}\left[\ln\left[\widehat{D}(q\lambda_m)/\Lambda\right]\right] - \mathbf{E}\left[\ln\left[\widehat{D}(\lambda_m)/\Lambda\right]\right]\right\}(2\ln q)^{-1}. \tag{.4.1}$$

Since the terms $\widehat{D}(q\lambda_m)$ and $\widehat{D}(\lambda_m)$ use the same number of frequencies, by Equation .2.3, $\mathbf{E}\left[\ln\left[\widehat{D}(q\lambda_m)/\Lambda\right]\right] = \mathbf{E}\left[\ln\left[\widehat{D}(\lambda_m)/\Lambda\right]\right] = \psi^{(0)}(m) +$

$\psi^{(0)}(m-1) - 2\ln(m)$. Therefore, $\mathbf{E}\left[\hat{b}_{AD}\right] = 0$.

2. Let L_1 and L_2 be identically distributed random variables with variance $\sigma_L^2 < \infty$ and correlation coefficient ρ_{L_1,L_2}, such that $|\rho_{L_1,L_2}| < 1$. Thus, $\mathbf{V}[L_1 - L_2] = 2\sigma_L^2(1-\rho_{L_1,L_2})$. Let now $L_1 \equiv \ln\left[\widehat{D}(q\lambda_m)/\Lambda\right](2\ln q)^{-1}$ and $L_2 \equiv \ln\left[\widehat{D}(\lambda_m)/\Lambda\right](2\ln q)^{-1}$. By the Equations .2.2 and .2.6, $\mathbf{V}[L_1] = \mathbf{V}[L_2] = \psi^{(1)}(m-1) + \psi^{(1)}(m)$. Therefore,

$$\mathbf{V}\ \left[\hat{b}_{AD}\right] = \left(\psi^{(1)}(m-1) + \psi^{(1)}(m)\right)(1-\rho_{L_1,L_2}) / \left\{2(\ln q)^2\right\}. \qquad (.4.2)$$

3. The proof of this part follows immediately from Lemma 3.

□

Capítulo 3

Additional Results: Robustness to outliers

The classical periodogram is a widespread tool in the context of the spectral analysis of time series. However, as pointed out by Fox (1972) it is very sensitive to the presence of outliers and, therefore, it becomes a useless tool in situations where the data is contaminated by atypical observations. Since additive outliers are quite common in practice, to define a new periodogram robust to the presence of these atypical observations is a valuable task that has a real practical interest.

Many approaches have been proposed in the time series literature in order to access robust periodograms see for example, the references discussed in Li (2008). Most of these references, however, concern weakly dependent time series.

In the long-memory framework, Molinares et al. (2009) suggested a robust plug-in periodogram, that is, a periodogram obtained by replacing, in the classical periodogram, the standard sample autocovariance by the robust autocovariance given in Ma and Genton (2000). In order to deal with outliers, Molinares et al.

(2009) introduced a robust estimator of the memory parameter d in the ARFIMA model. A review of robust methods in the frequency domain and their use in practical contexts can be found in Reisen and Molinares (2012). To verify the precision of the asymptotic theory on the robust estimation of the autocovariance and also on the parameters of the models, simulation studies for finite sample sizes were discussed in the previous references. In general, the empirical studies show that the estimators, based on the robust autocovariance, are outlier resistant and very useful in applied works. However, the robustness properties of the estimators weakens when the data deviates from the Gaussian assumption, that is, when the series has a heavy-tailed distribution.

This may be explained by the fact that the robust autocovariance function involves a constant which depends strongly on the Gaussian distribution assumption. In addition, the robust autocovariance estimator does not have the non-negative definite property, and becomes useless in the context of non-stationary process such as, for example, the ARFIMA process for $0.5 < d < 1.0$. This parameter range is quite often encountered in time series with long-memory (Hurvich and Ray (1995), Velasco (1999), Franco and Reisen (2007) among others).

To overcome some of the restrictive properties of the robust plug-in periodogram, a periodogram robust to the presence of outliers was proposed by Li (2010). Since the periodogram $I_X(\lambda)$ can be defined by $I_X(\lambda) = \frac{n}{4}\left\|\hat{\beta}_n(\lambda)\right\|^2$, where $\hat{\beta}_n(\lambda)$ is obtained by regressing series X_t against harmonic components $\mathbf{c_t}(\lambda) := [\cos(\lambda t), \sin(\lambda t)]^T$:

$$\hat{\beta}_n(\lambda) = \arg\min_{\beta\in\Re^2} \sum_{t=1}^{n} \left|X_t - \mathbf{c_t}(\lambda)^T\beta\right|^2 \qquad (3.0.1)$$

the robust periodogram is simply the solution of the above equation replacing the L_2 norm by an L_p norm:

$$\hat{\beta}_n(\lambda) = \arg\min_{\beta \in \Re^2} \sum_{t=1}^{n} \left| X_t - \mathbf{c_t}(\lambda)^T \beta \right|^p \qquad (3.0.2)$$

where $p \in (0, 2)$. If $p = 2$, then the ordinary periogram is obtained. Thus, the objective of this chapter is evaluate the performance of tests present in the previous chapter when the processes are contaminated by additive outliers and also the performance of the tests when the periodogram given in Equation 3.0.2 is considered when calculating the spectral density matrix. The empirical eveluation is divided in tow sets. The first set compares the GPH, AD and the LDR estimators performances when no outliers are included to the processes when additive outliers are present using the ordinary periodogram under the null hypothesis of non-cointegration. The second set analyzes the same estimators introducing the L_p norm for both: non-cointegrated and cointegrated series.

To the finite sample size investigation, 1500 processes with sample size $n = 250$ were generated for non cointegrated data, that is, $CI = (1, 1)$, and cointegrated data, that is, $CI = (1, 0)$. The parameter p varies in the set $\{1.75, 1.5, 1.2\}$. Innovations where generated as zero mean gaussian white noise processes with identity variance matrix. Each observation may contain an additive outlier in the error term of the cointegration equation (ε_t) with value equal 8 times the standard deviation with probability equal 0.01, that is, a random variable Ω is drawn for each point of the series such that, $\mathbb{P}(\Omega = 8) = \mathbb{P}(\Omega = -8) = 0.01$ and $\mathbb{P}(\Omega = 0) = 0.98$.

Table 3.1 presents the first set of simulations. For GPH two different specifications were considered. In the first one, GPH estimator was applied to OLS residuals using Santander et al. (2003) critical values. In the second specification, GPH was applied to differentiated OLS residuals and critical values are from $N(0,1)$. The results show that all considered estimators suffer significantly when the processes are contaminated by additive outliers. All tests produced biased estimates and consequently, remarkable size distortions. The worst case is the first difference GPH estimator. This test is completely distorted when the process is contaminated. The remainders estimators show quite similar results.

Since the all methods analyzed have showed size distortions, it can be worthwhile to replace the ordinary periodogram by the L_p periodogram in Equation 3.0.2. Now, table 3.2 displays the results for the mean of estimates, empirical standard deviations, mean squared error, size and power of the four considered tests. It can be seen that the GPH for the non-differentiated data is still oversized for the three values of parameter p. It was an expected result since the L_p periodogram cannot accommodate non-stationary data (for a wider discussion see Li (2010)). The remainders tests presented acceptable results when $p = \{1.5, 1.2\}$. All the three tests are still considerable oversized for $p = 1.75$. The best size performance was played by the LDR estimator and, although the AD estimator has displayed a better value than the LDR estimator when $p = 1.2$, the difference was small.

Evaluating the power, one can see that LDR estimator completely dominates the AD estimator but does not dominate the ΔGPH which has displayed the best results. However, the power performance of the three estimators (ΔGPH, AD and LDR) was very poor. An unexpected effect of the use of L_p periodogram was that the value of estimated parameter b has been showed very biased under the alternative hypothesis of non-cointegration. This suggests that L_p norm is not a very reliable option to estimate and test b. Additional research should investigate the reasons why this is occurring when the series are non-cointegrated even after differentiating. In addition, future research must evaluate the performance different robust peridograms (see Molinares et al. (2009) for examples).

Tabela 3.1: Estimates, size and power under outliers presence *vs* outliers free

Statistics	b_{GPH}[a]		$b_{\Delta GPH}$		b_{AD}[a]		b_{LDR}	
	No Outliers	Outliers	No Outliers	Outliers	No Outliers	Outliers	No Outliers	Outliers
Mean	0.9724	0.9119	-0.0202	-0.8454	0.0000	0.1456	0.0021	0.0862
sd	0.1035	0.1083	0.1020	0.1746	0.2790	0.2932	0.1389	0.1458
MSE	0.0115	0.0195	0.0108	0.7452	0.0776	0.1071	0.0193	0.0287
Rejection	0.0473	0.1293	0.0813	0.9973	0.0487	0.1373	0.0483	0.1387
σ_n	0.1000		0.1000		0.2795		0.1417	

[a] Since GPH was applied in the non-differentiated data, the expected value of b is 1.
[b] In order to compute σ_n to the AD estimator, values from Table 1 were used.

Tabela 3.2: Estimates, size and power using robust periodogram

		CI = (1,1)			CI = (1,0)		
Estimator	Statistics	p = 1.75	p = 1.5	p = 1.2	p = 1.75	p = 1.5	p = 1.2
GPH	Mean	0.8233	0.8001	0.7154	-0.0061	-0.0063	-0.0062
	sd	0.1034	0.1053	0.1168	0.0988	0.0985	0.0980
	mse	0.0419	0.0511	0.0946	0.0220	0.0224	0.0219
	Rejection	0.3447	0.4273	0.7360	1.0000	1.0000	1.0000
	σ_n		0.1000				
ΔGPH	Mean	-0.0450	-0.0229	-0.0091	-0.5931	-0.3778	-0.1993
	sd	0.1038	0.1003	0.0981	0.1113	0.1015	0.1003
	mse	0.0128	0.0106	0.0097	0.3641	0.1530	0.0498
	Rejection	0.1220	0.0807	0.0667	0.9993	0.9853	0.6133
	σ_n		0.1000				
AD	Mean	0.0832	0.0243	0.0204	0.8021	0.5160	0.2267
	sd	0.3130	0.3035	0.2934	0.3923	0.3430	0.3235
	mse	0.1048	0.0926	0.0864	0.7972	0.3839	0.1560
	Rejection	0.0960	0.0613	0.0567	0.7900	0.5320	0.2140
	σ_n		0.2972				
LDR	Mean	0.0404	0.0180	0.0155	0.5406	0.2905	0.1360
	sd	0.1446	0.1410	0.1420	0.1467	0.1413	0.1407
	mse	0.0225	0.0202	0.0204	0.2325	0.5233	0.7662
	Rejection	0.0853	0.0640	0.0607	0.9833	0.6320	0.2347
	σ_n		0.1417				

Capítulo 4

Conclusões

O presente trabalho investigou as propriedades do determinante da matriz de densidade espectral próximo à origem para um vetor bi-variado. O determinante é uma função potência do parâmetro que mensura a redução da ordem de integração da série de erros, b. A partir disto, dois estimadores foram propostos para tal parâmetro: o primeiro, baseado em Geweke and Porter-Hudak (1983), propôs uma regressão do logaritmo do determinante da matriz espectral do processo bivariado em estudo, o segundo um estimador semi-paramétrico do determinante médio baseado na proposta de Robinson (1994).

O artigo também propõe testes sob a hipótese nula de não cointegração derivados à partir dos estimadores apresentados. Estudos com amostras finitas foram realizados com o objetivo de avaliar, empiricamente, o desempenho dos estimadores e dos testes propostos através do cálculo do vício, do erro quadrático médio, dos níveis de significância e do poder. Os resultados apontam que os testes tem nível de significância empírico próximo do nível nominal. Além disso, o poder dos testes revelou um desempenho similar quando comparado com outros testes clássicos na literatura de cointegração discutidos em Dittmann (2000).

Os métodos aqui discutidos mostraram-se robustos a diferentes parametrizações da declividade da relação de cointegração (β). Foi investigada, ainda, as propriedades empíricas de tais métodos sob a presença de *outliers*. Neste sentido, o periodograma robusto à presença de outliers proposto em Li (2010) foi utilizado afim de obter-se testes resistentes à contaminação nos processos. O desempenho dos testes não foi satisfatório quando o periodograma L_p foi considerado. O estimadores apresentaram-se viciados para a hipótese alternativa e, consequentemente, o poder foi baixo. Pesquisas futuras devem investigar os motivos pelos quais isto ocorre. Neste sentido, o desempenho dos testes com outros periodogramas robustos à presença de *outliers* deve ser avaliado. Diversas alternativas são discutidas em Reisen and Molinares (2012) e uma sequência natural deste trabalho é avaliar o desempenho das mesmas.

Além da robustez a *outliers*, as pesquisas futuras investigarão o comportamento dos testes utilizando periodogramas robustos a dados faltantes bem como as propriedades assintóticas dos testes sob a hipótese de não-cointegração além das propriedades dos testes sob o relaxamento de algumas hipóteses aqui assumidas como: erros ruídos brancos e séries integradas de ordem 1. Por fim, uma versão multivariada destes testes deve ser considerada com o intuito de testar múltiplos vetores de cointegração.

Bibliography

Abramowitz, M. and I. A. Stegun (Eds.) (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Volume 55 of *National Bureau of Standards Applied Mathematics Series.* Washington, DC: United States Government Printing Office.

Aloy, M., M. Boutahar, K. Gente, and A. Péguin-Feissolle (2013). Long-run relationships between international stock prices: further evidence from fractional cointegration tests. *Applied Economics 45*(7), 817–828.

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

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1987). *Regular Variation*, Volume 27 of *Encyclopedia of Mathematics and its Applications.* Cambridge: Cambridge University Press.

Brillinger, D. R. (1981). *Time series: data analysis and theory*, Volume 36 of *Classics in Applied Mathematics.* Philadelphia: Siam.

Cai, T. T., T. Liang, and H. H. Zhou (2013). Law of log determinant of sample covariance matrix and optimal estimation of differential entropy for high-dimensional gaussian distributions. *arXiv preprint arXiv:1309.0482*.

Cheung, Y. W. and K. S. Lai (1993). A fractional cointegration analysis of purchasing power parity. *Journal of Business & Economic Statistics 11*(1), 103–112.

Chung, K. L. (2001). *A course in probability theory*. San Diego: Academic Press.

Cuestas, J. C., L. A. Gil-Alana, and K. Staehr (2014). Government debt dynamics and the global financial crisis: Has anything changed in the EA12? *Economics Letters 124*(1), 64–66.

Dittmann, I. (2000). Residual-based tests for fractional cointegration: A monte carlo study. *Journal of Time Series Analysis 21*(6), 615–647.

Dittmann, I. (2001). Fractional cointegration of voting and non-voting shares. *Applied Financial Economics 11*(3), 321–332.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society 55*(2), 251–276.

Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society 34*, 350–363.

Franco, G. C. and V. A. Reisen (2007). Bootstrap approaches and confidence intervals for stationary and non-stationary long-range dependence processes. *Physica A: Statistical Mechanics and its Applications 375*(2), 546–562.

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

Goodman, N. (1963). The distribution of the determinant of a complex wishart distributed matrix. *The Annals of Mathematical Statistics 34*(1), 178–180.

Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *Journal of econometrics 16*(1), 121–130.

Hamilton, J. D. (1994). *Time series analysis*, Volume 2. Princeton: Princeton University Press.

Hosking, J. R. (1981). Fractional differencing. *Biometrika 68*(1), 165–176.

Hurvich, C. M. and K. I. Beltrao (1994). Automatic semiparametric estimation of the memory parameter of a long-memory time series. *Journal of Time Series Analysis 15*(3), 285–302.

Hurvich, C. M. and B. K. Ray (1995). Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes. *Journal of time series analysis 16*(1), 17–41.

Jeffrey, A. and D. Zwillinger (Eds.) (2007). *Table of integrals, series, and products*. San Diego: Academic Press.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: Journal of the Econometric Society 59*(6), 1551–1580.

Kanas, A. (1998). Linkages between the US and European equity markets: further evidence from cointegration tests. *Applied Financial Economics 8*(6), 607–614.

Li, T.-H. (2008). Laplace periodogram for time series analysis. *Journal of the American Statistical Association 103*(482), 757–768.

Li, T.-H. (2010). A nonlinear method for robust spectral analysis. *Signal Processing, IEEE Transactions on 58*(5), 2466–2474.

Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica 59*(5), 1279–1313.

Lobato, I. N. and P. M. Robinson (1998). A nonparametric test for I(0). *The Review of Economic Studies 65*(3), 475–495.

Ma, Y. and M. G. Genton (2000). Highly robust estimation of the autocovariance function. *Journal of time series analysis 21*(6), 663–684.

MacKinnon, J. G. (1991). Critical values for cointegration tests. In R. F. Engle and C. W. J. Granger (Eds.), *Long-Run Economic Relationships: Readings in Cointegration*, Chapter 13. Oxford: Oxford University Press.

Marinucci, D. and P. M. Robinson (2001). Semiparametric fractional cointegration analysis. *Journal of Econometrics 105*(1), 225–247.

McHale, I. and D. Peel (2010). Habit and long memory in UK lottery sales. *Economics Letters 109*(1), 7–10.

Molinares, F. F., V. A. Reisen, and F. Cribari-Neto (2009). Robust estimation in long-memory processes under additive outliers. *Journal of Statistical Planning and Inference 139*(8), 2511–2525.

Nielsen, M. Ø. (2004). Spectral analysis of fractionally cointegrated systems. *Economics Letters 83*(2), 225–231.

Phillips, P. C. and S. Ouliaris (1988). Testing for cointegration using principal components methods. *Journal of Economic Dynamics and Control 12*(2), 205–230.

Priestley, M. B. (1981). *Spectral analysis and time series.* London: Academic Press.

Reisen, V. A. and F. F. Molinares (2012). Robust estimation in time series with long and short memory properties. In *Annales Mathematicae et Informaticae*, Volume 39, pp. 207–224.

Robinson, P. M. (1994). Semiparametric analysis of long-memory time series. *The Annals of Statistics 22*(1), 515–539.

Robinson, P. M. and D. Marinucci (1998). Semiparametric frequency domain analysis of fractional cointegration. STICERD Econometrics Discussion Paper 348, London School of Economics.

Robinson, P. M. and D. Marinucci (2001). Narrow-band analysis of nonstationary processes. *Annals of Statistics 29*(4), 947–986.

Robinson, P. M. and Y. Yajima (2002). Determination of cointegrating rank in fractional systems. *Journal of Econometrics 106*(2), 217–241.

Santander, L. A. M., V. A. Reisen, and B. Abraham (2003). Non-cointegration tests and fractional arfima process. *Statistical Methods 5*(1), 1–22.

Velasco, C. (1999). Non-stationary log-periodogram regression. *Journal of Econometrics 91*(2), 325–371.

Velasco, C. (2003). Nonparametric frequency domain analysis of nonstationary multivariate time series. *Journal of Statistical Planning and Inference 116*(1), 209–247.