https://doi.org/10.26219/heal.aueb.5789

**ATHENS UNIVERSITY
OF ECONOMICS AND BUSINESS**

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

Spectral Analysis of Time Series of Interest Rates

By

Sofia Antoniou Mpatzavali

A THESIS

Submitted to the Department of Statistics

of the Athens University of Economics and Business

in partial fulfilment of the requirements for

the degree of Master of Science in Statistics

Athens, Greece

October 2007

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΤΙΤΛΟΣ

Φασματική ανάλυση χρονολογικών σειρών επιτοκίων

Σοφία Αντωνίου Μπατζάβαλη

ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα
Οκτώβριος 2007

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

DEDICATION

Dedicated to the one who made me believe in what I can do.

https://doi.org/10.26219/heal.aueb.5789

ACKNOWLEDGEMENTS

To my supervisor, whose advises and guidance were valuable for the implementation of this thesis. I will also thank my fiancé for his valuable help.

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

VITA

STUDIES:

2003-. Athens University of Economics and Business, Postgraduate in Department of Statistics

1997-2003 University of Patras, Department of Exact Sciences, Degree in Mathematics

PROFESSIONAL EXPERIENCE:

12/06- OTE, Sub direction of National Providers Commercial Operation,:

i. Product Manager of Physical Co-location, and Remote Location of National Providers
ii. WholeSale CRM administrator
iii. Reporting of Directions Operations

09/05-11/06 OTE, Department of Division, Strategic Development of Retail Network Division, position: SPSS Specialist

i. Setting up the procedures in order to obtain the ISO 9001 Certification
ii. CRM Developing Consultant
iii Annual Sales Forecasting
iv Weekly and monthly Sales Reporting
v Forecasting of employee requirements, in order to announce job vacancies

2005 Tzaneio Hospital, Department of Data Processing and Statistics, position: Publication of statistical reports

2000-2001 West Coast Food Corporation, position: Assistant Manager of 'La Pasteria' branch at Patras

FOREIGN LANGUAGES:

English: Very good knowledge a) First Certificate in English, Cambridge University b) holder of Toefl degree

French: Basic knowledge (Francoschool)

COMPUTER KNOWLEDGE:

2005 Certificate of K.E.K ΣΒΙΕ, "Microsoft Visual Basic Software Developers", Operational Program "Employment and Professional Training", Ministry of Employment and Social Protection with the co financing of European Union, (400 hours).

2003-2005 Operation of Statistical programs: Spss, S-Plus, E-views, Minitab

2000 Certificate of Mathematician Society in Windows, Ms-Office: Word, Excel, Access, PowerPoint, Internet, Outlook Express, Ms-Dos (400 hours)

https://doi.org/10.26219/heal.aueb.5789

WORKS-STUDIES:

1) Post-graduate Work: Forecasting and Analysis of VAR stock market models with Spectral Analysis and Cointegration methods
2) Graduate work in Mathematical Programming, subject: Distributive Algorithm in Linear Programming

CONGRESSES:

2004 Attendance of 17th Pan-Hellenic Statistical Congress

INTERESTS:

Swimming, cinema, music, literature

DISTINCTIONS:

1995 Gold medal of Pan-Hellenic competition in literature

https://doi.org/10.26219/heal.aueb.5789

ABSTRACT

Sofia Mpatzavali

Spectral Analysis of Time Series of Interest Rates

October 2007

Cointegration test is applied to a set of interest rates of five countries for determining whether there exists any long term relationship between the series. Spectral analysis, a frequency based technique, is applied to the same set of series in order to examine whether there are any correlations among their frequency-λ-components, in particular for low λ. Such correlations would also indicate a certain degree of long-term co-movement of the series. Spectrum, coherency and phase are interpreted for this purpose. Non-parametric, as well as parametric (via a VAR fit) estimators, of these quantities are presented and compared to each other.

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

ΠΕΡΙΛΗΨΗ

Σοφία Μπατζάβαλη

Φασματική ανάλυση χρονολογικών σειρών επιτοκίων

Οκτώβριος 2007

Σε ομάδα επιτοκίων πέντε χωρών εφαρμόζουμε cointegration test αποσκοπώντας να διερευνήσουμε αν υπάρχουν μακροπρόθεσμες συσχετίσεις μεταξύ των χρονοσειρών τους. Η φασματική ανάλυση, μέθοδος που βασίζεται σε παρατήρηση συχνοτήτων, εφαρμόζεται στην ίδια ομάδα χωρών διερευνώντας κατά πόσο υπάρχουν συσχετίσεις μεταξύ των λ-συχνοτήτων τους, ειδικά των χαμηλών. Τέτοιες συσχετίσεις θα υποδείκνυαν και ένα συγκεκριμένο βαθμό μακροπρόθεσμης επίδρασης. Το φάσμα, η συνάφεια, και η φάση εξετάζονται γι' αυτό τον σκοπό. Μη παραμετρικοί, όπως και παραμετρικοί εκτιμητές (μέσω ενός πολυμεταβλητού μοντέλου) αυτών των ποσοτήτων, παρουσιάζονται και συγκρίνονται μεταξύ τους.

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

TABLE OF CONTENTS

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

LIST OF TABLES

Table **Page**

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

LIST OF FIGURES

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

CHAPTER 1

INTRODUCTION

1.1 Introduction

Time series techniques are widely used in the analysis of financial data and for estimating and testing models for asset prices and returns like the capital asset pricing model and the arbitrage pricing model. They are used to uncover and exploit predictive relationships between financial variables. Time series techniques are also used for testing the informational efficiency of financial markets.

The present thesis is motivated by Juselius (2001) paper 'Big Shocks, Outliers, and Interventions. A Cointegration and Common Trends Analysis of Daily Bond Rates'. Juselius worked on co-movements and pushing forces in nine international daily bond rate data. Economic theory often implies equilibrium relationships between the levels of time series variables that are best described as being $I(1)$. Similarly, arbitrage arguments imply that the $I(1)$ prices of certain financial time series are linked. Cointegration between non stationary time series exists if there is a linear combination of them that is stationary. This linear combination is referred to as a long-run equilibrium relationship.

Our purpose in this thesis is to explore whether similar conclusions, as those which Juselius derived from daily bond rate data, could be drawn for monthly interest rates of a subset of the same countries.

Many time series arising in practice are best considered as components of some vector-valued (multivariate) time series $\{\mathbf{X}_t\}$ whose specification includes not only the serial dependences of its component series $\{X_{ti}\}$ but also the interdependence between different component series $\{X_{ti}\}$ and $\{X_{tj}\}$. The spectral representation of a stationary process $\{\mathbf{X}_t, t = 0, \pm 1, ...\}$ essentially decomposes $\{\mathbf{X}_t\}$ into a sum of sinusoidal components with uncorrelated random coefficients. In conjunction with this decomposition there is a corresponding decomposition into sinusoids of the autocovariance function of $\{\mathbf{X}_t\}$. This decomposition is then represented by the spectrum of each series, the coherency and phase spectrum for each pair of series, which all of them are real functions of $(-\pi,\pi)$. The spectrum of a series at frequency λ provides information on the strength of the presence of frequency λ components in the series. The coherency of a pair of series at frequency λ represents the correlation between frequency λ components of the two series. The phase of a pair at λ gives the phase lag between λ component of the two series. The spectral decomposition is thus an analogue for stationary stochastic processes of the more familiar Fourier representation of deterministic functions. The analysis of stationary processes by means of their spectral representations is often referred to as the "frequency domain" analysis of time series. It is equivalent to "time domain" analysis, based on the autocovariance function, but provides an alternative way of viewing the process which for some applications may be more illuminating. In order to apply multivariate spectral analysis in a set of series it is essential that the series are stationary.

In order to visualize the usage of these techniques we are going to implement them for a five dimensional system consisting of the monthly interest rates of countries: **France, United Kingdom, Netherlands, Denmark and United States.** The data are from 1973 to 1998 in monthly frequency but we will only use data starting from 1974 for the analysis.

The original data are descriptively analyzed and found to have linear trends. Moreover it will turn out that these monthly data are unit root as shown in section *5.4.2 Unit Root*, so the spectral analysis is conducted with the differences of the series because the levels of the series are non-stationary. Also we apply a cointegration test in *4.3 Cointegration Analysis of Interest Rates*. Afterwards we will use frequency domain analysis to investigate whether there are any strong relations between long-medium term components of data. We will present the spectrum of the spectral analysis model together with the spectrum estimation occurred from the VAR model whose order is determined by the AIC value (*4.4.1 Individual spectra*). Next, in *5.4.2 Coherency-Phase between Interest Rates,* we oppose the Coherency and the Phase diagrams of each pair of series in order to view the magnitude of their correlation along frequency.

https://doi.org/10.26219/heal.aueb.5789

CHAPTER 2

MULTIVARIATE TIME SERIES

2.1 Multivariate ARMA process

Many time series arising in practice are best considered as components of some vector valued (multivariate) time series $\{\mathbf{X}_t\}$:=(X_{t1}, X_{t2}, ….,X_{tm})′ t=0,±1,±2,... whose specifications include not only the serial dependence of each component, but also the interdependence between different component series $\{X_{ti}\}$ and $\{X_{tj}\}$, in particular, a stationary time series is specified by its mean vector and its covariance matrices. Below we introduce the ARMA process, its mean vector, covariance matrix and then predict the multivariate series process.

Consider m time series {X_{ti} t=0,±1, ±2, ±3…}, i= 1….. m with $EX_{ti}^{\ 2}<\infty$ for all t and all i. If all the finite dimensional joint distributions of the random variables {X_{ti}} were multivariate normal, then the distributional properties of {X_{ti}} would be completely determined by the means,

$\mu_{ti} := EX_{ti},$

and covariances,

$\gamma_{ij}(t+h) := E[(X_{t+h,i},-\mu_{t+h,i})(X_{t,j},-\mu_{t,j})]$.

Even when the observations {X_{ti}} do not have joint normal distributions, the quantities μ_{ti} and $\gamma_{ij}(t+h,t)$ specify the second order properties, the covariances providing us with a measure of dependence among observations in the same (autocovariance) or in different series (cross correlation for i≠j.

It is more convenient when dealing with m interrelated series to use vector notation as above.

We define the mean vectors,

$\boldsymbol{\mu}_t:=E\mathbf{X}_t=(\mu_{t1,\ldots,}\mu_{tm})'$

and covariance matrices,

$\Gamma(t+h,t):=E[(\mathbf{X}_{t+h}-\boldsymbol{\mu}_{t+h})(\mathbf{X}_t-\boldsymbol{\mu}_t)^*]$,

where * denotes complex conjugate transpose. However we shall assume that X_t is real.

https://doi.org/10.26219/heal.aueb.5789

Def.2.1 (*Stationary Multivariate Time Series*)

The series $\{\mathbf{X}_t\}$ t=0,±1, ±2,…, is said to be stationary if $\boldsymbol{\mu}_t$ and $\Gamma(t+h,t)$, h=0,±1,…, are independent of t.

For a stationary time series we shall use the notation,

$\boldsymbol{\mu}:= E\mathbf{X}_t=(\mu_1,\ldots,\mu_m)'$,

and

$\Gamma(h):=E[(\mathbf{X}_{t+h}-\boldsymbol{\mu})(\mathbf{X}_t-\boldsymbol{\mu})']=[\gamma_{ij}(h)]_{i,j=1}^{m}$.

We shall refer to $\boldsymbol{\mu}$ as the mean of the series and $\Gamma(h)$ as the ***covariance matrix*** at lag h, and to γ_{ii} as the autocovariance function for each i. The function $\gamma_{ij}(\cdot), i\neq j$, is called the cross-covariance function of the two series $\{X_{ti}\}$, $\{X_{tj}\}$. It should be noted that γ_{ij} is not in general the same as γ_{ji} The ***correlation matrix function*** $R(\cdot)$ is defined by

$$R(h):=[\frac{\gamma_{ij}(h)}{[\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}}]^{m}_{i,j=1}=[\rho_{ij}(h)]_{i,j=1}^{m}$$

The function $R(\cdot)$ is the covariance matrix function of the normalized series obtained by subtracting $\boldsymbol{\mu}$ from $\mathbf{X}_t$ and then dividing each component by its standard deviation.

The covariance matrix function $\Gamma(\cdot)=[\gamma_{ij}(\cdot)]_{i,j=1}^{m}$, of a stationary time series {Xt} has the properties,

$\Gamma(h)=\Gamma(-h)$,

$|\gamma_{ij}(h)|\leq[\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}$, i,j=1,2,…m

$\gamma_{ii}(\cdot)$ is an autocovariance function, i=1,2,…m

$\sum_{j,k=1}^{n} a'_j\Gamma(j-k)a_k\geq 0$ for all n϶ {1,2,….} and $a_1, a_2,\ldots,a_n \in R^m$.

Properties (i), (ii), (iii) and (iv) are shared by the correlation matrix function $R(\cdot)=[\rho_{ij}(\cdot)]^{m}_{i,j=1}$ which has the additional property,

v) $\rho_{ii}(0)=1$.

Def.2.2 (*Multivariate ARMA(p,q)* Process) $\{\mathbf{X}_t, t=0, \pm1, \pm2, \ldots\}$ is an m-variate ARMA(p,q) process if $\{\mathbf{X}_t\}$ is a stationary solution of the difference equations,

$\mathbf{X}_t - \mathbf{\Phi}_1\mathbf{X}_{t-1} - \cdots - \mathbf{\Phi}_p\mathbf{X}_{t-p} = \mathbf{Z}_t + \mathbf{\Theta}_1\mathbf{Z}_{t-1} + \cdots \mathbf{\Theta}_q\mathbf{Z}_{t-q}$ or more compactly

$\mathbf{\Phi}(\mathbf{B})\mathbf{X}_t = \mathbf{\Theta}(\mathbf{B})\mathbf{Z}_t$

where $\mathbf{\Phi}_1, \cdots, \mathbf{\Phi}_p, \mathbf{\Theta}_1, \cdots, \mathbf{\Theta}_q$ are real $m \times m$ matrices and $\{\mathbf{Z}_t\} \sim WN(0, \Sigma)$

and

$\mathbf{\Phi}(z) = \mathbf{I} - \mathbf{\Phi}_1 z - \cdots - \mathbf{\Phi}_p z^p$

$\mathbf{\Theta}(z) = \mathbf{I} + \mathbf{\Theta}_1 z + \cdots + \mathbf{\Theta}_q z^q$

are matrix valued polynomials, I is the $m \times m$ identity matrix and B denotes the backward shift operator. Each component of the matrices $\mathbf{\Phi}(z)$, $\mathbf{\Theta}(z)$ is a polynomial with real coefficients and degree less than or equal to p, q respectively.

It is expedient to express the covariance matrix $\Gamma(h) = E(\mathbf{X}_{t+h}\mathbf{X}'_t)$ of the causal Multivariate ARMA(p,q) process as

$\Gamma(h) = \sum_{k=0}^{\infty} \Psi_{h+k}\Sigma\Psi'_{\kappa}$ h=0, ±1, ±2,…., where the matrices Ψ_j are found from

$\Psi(z) := \sum_{j=0}^{\infty} \Psi_j z^j = \Phi^{-1}(z)\Theta(z)$, $|z| \le 1$

and $\Psi_j := 0$ for j<0.The covariance matrices $\Gamma(h)$, h=0, ±1, ±2,......., can be determined by solving the Yule Walker equations,

$\Gamma(j) - \sum_{r=1}^{p} \Phi_r \Gamma(j-r) = \sum_{j \le r \le q} \Theta_r \Sigma \Psi'_{r-j}$, j=0,1,2,

obtained by post-multiplying $\mathbf{X}_t - \mathbf{\Phi}_1\mathbf{X}_{t-1} - \cdots - \mathbf{\Phi}_p\mathbf{X}_{t-p} = \mathbf{Z}_t + \mathbf{\Theta}_1\mathbf{Z}_{t-1} + \cdots \mathbf{\Theta}_q\mathbf{Z}_{t-q}$ by X_{t-j}' and taking expectations.

In this work we are primarily interested in Multivariate ARMA(p,0) (or Var(p)) process i.e. a ***Multivariate AR(p)*** which satisfies the equation

$\mathbf{\Phi}(B)\mathbf{X}_t = \mathbf{Z}_t$ (1.1)

where Z_t is m-variate White Noise sequence $Z_t = (Z_{1,t}\ Z_{2,t}, Z_{3,t}, \ldots, Z_{m,t})^T$ with covariance matrix $E(Z_t Z_t^T) = \Sigma$ and $EZ_t Z_s = 0, t \neq s$, which is assumed to be nonsingular, $\Phi(z)$ is a $m \times m$ lag polynomial matrix with elements $\varphi_{ij}(z)$. We further assume that $\Phi(0)=I_m$ the identity matrix and that $\det(\Phi(z))\neq 0$ for $|z|\leq 1$. This latter condition is necessary and sufficient for the existence of a stationary and causal solution.

2.2 Best linear predictors for AR process

In this section we shall confine our discussion to bivariate time series but the ideas, as *Brockwell & Davis (1991)* mention, can easily be extended to higher dimensions and to series whose covariances are not absolutely summable.

Suppose we have observations $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n$ of a zero –mean stationary *m*-variate time series and let $\hat{\Gamma}(o), \ldots, \hat{\Gamma}(p-1)$ be the sample covariance function estimates. Then the fitted AR(p) process (p<n) is

$$\mathbf{X}_t = \hat{\mathbf{\Phi}}_{p1}\mathbf{X}_{t-1} + \ldots + \hat{\mathbf{\Phi}}_{pp}\mathbf{X}_{t-p} + \mathbf{Z}_t \quad \{\mathbf{Z}_t\} \sim WN(\mathbf{0}, \hat{\mathbf{V}}_p),$$

where coefficients $\hat{\mathbf{\Phi}}_{p1}, \ldots, \hat{\mathbf{\Phi}}_{pp}$ and $\hat{\mathbf{V}}_p$ are computed recursively from :

$$\begin{aligned}
\mathbf{\Phi}_{pp} &= \mathbf{\Delta}_{p-1}\hat{\mathbf{V}}_{p-1}^{-1}, \\
\tilde{\mathbf{\Phi}}_{pp} &= \hat{\mathbf{\Delta}}_{p-1}\mathbf{V}^{-1}{}_{p-1}, \\
\mathbf{\Phi}_{pk} &= \mathbf{\Phi}_{p-1,k} - \mathbf{\Phi}_{pp}\hat{\mathbf{\Phi}}_{p-1,p-k}, \quad k = 1, \ldots.., p-1 \\
\hat{\mathbf{\Phi}}_{pk} &= \hat{\mathbf{\Phi}}_{p-1,k} - \hat{\mathbf{\Phi}}_{pp}\mathbf{\Phi}_{p-1,p-k}, \quad k = 1, \ldots.., p-1
\end{aligned}$$

and $\mathbf{V}_p, \hat{\mathbf{V}}_p, \mathbf{\Delta}_p, \hat{\mathbf{\Delta}}_p$ are given by

$$\begin{aligned}
\mathbf{V}_p &= \hat{\mathbf{\Gamma}}(0) - \mathbf{\Phi}_{p1}\hat{\mathbf{\Gamma}}(-1) - \ldots.. - \mathbf{\Phi}_{pp}\hat{\mathbf{\Gamma}}(-p) \\
\hat{\mathbf{V}}_p &= \hat{\mathbf{\Gamma}}(0) - \hat{\mathbf{\Phi}}_{p1}\hat{\mathbf{\Gamma}}(-1) - \ldots - \hat{\mathbf{\Phi}}_{pp}\hat{\mathbf{\Gamma}}(-p) \\
\mathbf{\Delta}_p &= \hat{\mathbf{\Gamma}}(p+1) - \mathbf{\Phi}_{p1}\hat{\mathbf{\Gamma}}(p) - \ldots - \mathbf{\Phi}_{pp}\hat{\mathbf{\Gamma}}(-1) \\
\hat{\mathbf{\Delta}}_p &= \hat{\mathbf{\Gamma}}(-p-1) - \hat{\mathbf{\Phi}}_{p1}\hat{\mathbf{\Gamma}}(-p) - \ldots - \hat{\mathbf{\Phi}}_{pp}\hat{\mathbf{\Gamma}}(-1)
\end{aligned}$$

where $\begin{matrix} V_0 = \hat{V}_0 = \hat{\Gamma}(0) \\ \Delta_0 = \hat{\Delta}_0 = \hat{\Gamma}(1) \end{matrix}$ and n=1,…,p

The order of the auto regression may be chosen to minimize

$$AIC = -2\ln L(\hat{\mathbf{\Phi}}_{p1}, \ldots, \hat{\mathbf{\Phi}}_{pp}, \hat{\mathbf{V}}_p) + 2(pm^2+1)nm/(nm - pm^2 - 2).$$

https://doi.org/10.26219/heal.aueb.5789

Generally modeling with financial time series requires some care because the time series properties of the data can influence the properties of standard estimates and inference methods. In general, standard regression techniques are appropriate for analysis of $I(0)$ stationary and non-trending data. In the next chapter we will discuss the handling of such data.

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

CHAPTER 3

SPECTRAL REPRESENTATION AND ANALYSIS OF TIME SERIES

3.1 Basic components for spectral representation

In this chapter we will assume for simplicity that we deal with bivariate time series (m=2) and all the conclusions could stand for every multivariate case. As *Hamilton James D.* (1994) apposes:

Def.3.1 (*The Cross Spectrum*) If $\{\mathbf{X}_t\}=\{X_{t1},X_{t2}\}$ is a stationary bivariate time series with mean **0** and covariance matrix function $\Gamma(\cdot)$ satisfying

$$\sum_{-\infty}^{\infty}\left|\gamma_{ij}(h)\right|<\infty,\ i,j=1,2$$

then the matrix

$$f(\lambda)=\frac{1}{2\pi}\sum_{-\infty}^{\infty}e^{-ih\lambda}\Gamma(h)=\begin{pmatrix} f_{11}(h) & f_{12}(h) \\ f_{21}(h) & f_{22}(h)\end{pmatrix}$$

is called the spectral density matrix or spectrum of $\{X_t\}$.

where the function $f_{12}(\lambda)=\frac{1}{2\pi}\sum_{-\infty}^{\infty}e^{-ih\lambda}\gamma_{12}(h)$ $\lambda\in[-\pi,\pi]$ is called the cross spectral density of $\{X_{t1}\}$ and $\{X_{t2}\}$.

The spectral representations of $\gamma_{ij}(h)$ and $\Gamma(h)$ follow from this definition. It can be shown that:

$$\gamma_{ij}(h)=\int_{-\pi}^{\pi}e^{ih\lambda}f_{ij}(\lambda)d\lambda,\ i,j=1,2$$

and

$$\Gamma(h)=\int_{-\pi}^{\pi}e^{ih\lambda}f(\lambda)d\lambda\ \ i,j=1,2$$

The function $f_{ii}(\cdot)$ is the spectral density of the univariate series $\{\mathbf{X}_{ti}\}$ and is therefore real valued and symmetric about zero. However since $\gamma_{ij}(\cdot),\ i\neq j$, is not in general symmetric about zero, the cross spectrum $f_{ij}(\cdot)$ is typically complex valued.

 https://doi.org/10.26219/heal.aueb.5789

In the multivariate case there exists a spectral representation of the series itself, similar to the one in the univariate case. As far concerns a univariate series $\{X_t\}$, we know that there exists an orthogonal increment process $\{[Z_i(\lambda), -\pi \le \lambda \le \pi\}$ in the spectral representation of the univariate series, where

$$X_{ti} = \int_{(-\pi,\pi)} e^{it\lambda} dZ_i(\lambda)$$

and $f_{ii}(\lambda)d\lambda = E|dZ_i(\lambda)|^2$,

the latter being an abbreviation for $\int_{\lambda_1}^{\lambda_2} f_{ii}(\lambda)d\lambda = E|Z_i(\lambda_2) - Z_i(\lambda_1)|^2$, $-\pi \le \lambda_1 \le \lambda_2 \le \pi$. In the bivariate case we have:

Theorem 3.1

There exists an orthogonal increment process $Z_t = (Z_{1,t}, Z_{2,t})$ with property $E(dZ_i(\lambda)\overline{dZ_j(\mu)}) = 0$ for $\lambda \ne \mu$ and i,j=1,2, such that

$$f_{ij}(\lambda)d\lambda = E(dZ_i(\lambda)\overline{dZ_j(\lambda)}) \tag{3.1}$$

which is shorthand for

$$\int_{\lambda_1}^{\lambda_2} f_{ii}(\lambda)d\lambda = E\left[\left(Z_i(\lambda_2) - Z_i(\lambda_1)\right)\left(\overline{Z_j(\lambda_2)} - \overline{Z_j(\lambda_1)}\right)\right], \ -\pi \le \lambda_1 \le \lambda_2 \le \pi .$$

In order to interpret the cross spectrum we introduce the coherency.

Def.3.2 (*Coherency)* K_{12}, which is calculated by $K_{12}(\lambda) = f_{12}(\lambda)/[f_{11}(\lambda)f_{22}(\lambda)]^{1/2}$ is called coherency or coherence measure and is the interpretation of the correlation between $dZ_1(\lambda)$ and $dZ_2(\lambda)$.

By the Cauchy Schwartz inequality, the squared coherency function $|K_{12}(\lambda)|^2$ satisfies the inequalities,

$$0 \le |K_{12}(\lambda)|^2 \le 1, \quad -\pi \le \lambda \le \pi,$$

and a value near one indicates a strong relationship between $dZ_1(\lambda)$ and $dZ_2(\lambda)$.

Since $f_{12}(\lambda)$ is complex-valued, it can be expressed as

$$f_{12}(\lambda) = c_{12}(\lambda) - iq_{12}(\lambda),$$

where $c_{12}(\lambda) = \mathrm{Re}\{f_{12}(\lambda)\}$

and $q_{12}(\lambda)=-\mathrm{Im}\{f_{12}(\lambda)\}$.

The function $c_{12}(\lambda)$ is called cospectrum of $\{X_{t1}\}$ and $\{X_{t2}\}$, and $q_{12}(\lambda)$ is called the quadrature spectrum.

Alternatively $f_{12}(\lambda)$ can be expressed in polar coordinates as

$$f_{12}(\lambda)=\alpha_{12}(\lambda)\exp[i\varphi_{12}(\lambda)],$$

where

$$\alpha_{12}(\lambda)=(c_{12}(\lambda)^2+ q_{12}(\lambda)^2)^{1/2}$$

is called amplitude spectrum and

$$\varphi_{12}(\lambda)=\arg(c_{12}(\lambda)-i\, q_{12}(\lambda))\in(-\pi,\pi],$$

is called the phase spectrum of $\{X_{t1}\}$ and $\{X_{t2}\}$. The coherency is related to the phase and amplitude spectra by:

$$K_{12}(l)= a_{12}(l)[f_{11}(l)f_{22}(l)]^{-1/2}\exp[if_{12}(l)]=|K_{12}(l)|\exp[if_{12}(l)]$$

Remark

Coherency has some special properties according to the relation of the series whose correlation it depicts.

Proposition 1

If $\{X_{t1}\}$, $\{X_{t2}\}$ are related by a time invariant linear filter, then $|K_{12}(\lambda)|^2=1,\ -\pi\le\lambda\le\pi$

Proposition 2

If $\{X_{t1}\}$, $\{X_{t2}\}$ have squared coherency $|K_{12}(\lambda)|^2$ and if linear filters are applied to each process giving

$$Y_{t1}=\sum_{j=-\infty}^{j=\infty} a_j X_{t-j,1}$$

and

$$Y_{t2}=\sum_{j=-\infty}^{j=\infty} \beta_j X_{t-j,2}$$

 https://doi.org/10.26219/heal.aueb.5789

where $\sum_j |a_j| < \infty$ and $\sum_j |\beta_j| < \infty$ then $\{Y_{t1}\}$ and $\{Y_{t2}\}$ have the same squared coherency $|K_{12}(\lambda)|^2$.

Let $\{\mathbf{X}_t\}$ be a bivariate stationary series and consider the problem of finding a time-invariant linear filter $\mathbf{\Psi} = \{\psi_j\}$ which minimizes $E\left|X_{t2} - \sum_{j=-\infty}^{j=\infty} \psi_j X_{t-j,1}\right|^2$. It can be shown that this $\mathbf{\Psi}$ is $\psi(e^{-iv}) = \dfrac{\mathrm{E}(dZ_2(v)\overline{dZ_1(v)})}{\mathrm{E}|dZ_1(v)|^2} = f_{21}(v)/f_{11}(v)$, and the spectral density of $\sum_j \psi_j X_{t-j,1}$ is then $f_{2/1}(v) = |f_{21}(v)|^2 / f_{11}(v)$. The density $f_{2/1}$ is thus the spectral density of the linearly filtered version of $\{X_{t1}\}$, which is the best mean square approximation to $\{X_{t2}\}$. We also observe that $|K_{12}(\lambda)|^2 = \dfrac{f_{2/1}(v)}{f_{22}(v)}$, so that $|K_{12}(\lambda)|^2$ can be interpreted as the proportion of the variance of $\{X_{t2}\}$, at frequency v which can be attributed to a linear relationship between $\{X_{t2}\}$ and $\{X_{t1}\}$.

If $\{X_{t1}\}$ and $\{X_{t2}\}$ are uncorrelated, then $f_{12}(v) = 0, -\pi \le v \le \pi$, from which it follows that $|K_{12}(\lambda)|^2 = 0,\ -\pi \le v \le \pi$.

Next we consider problems of statistical inference for time series based on frequency-domain properties of the series. The fundamental tool used is the periodogram.

3.2 The periodogram

Consider an arbitrary set of observations $x_1, \ldots, x_n \in \mathbb{C}$ made at times 1,…,n respectively. The vector

$$\mathbf{x} := (x_1, \ldots, x_n)', \tag{3.2}$$

belongs to the n-dimensional complex space $\mathbb{C}^n$. If u and v are two elements of $\mathbb{C}^n$, we define this inner product of u and v:

$$\langle u, v \rangle = \sum_{i=1}^{n} u_i \bar{v}_i.$$

https://doi.org/10.26219/heal.aueb.5789

Each x_t can be expressed as a linear combination of harmonics,

$$x_t = n^{-1/2} \sum_{-\pi<\omega_j\le\pi} a_j e^{it\omega_j}, t = 1,\dots,n, \tag{3.3}$$

where the frequencies $\omega_j=2\pi j/n$ are the integer multiples of the fundamental frequency $2\pi/n$ which fall in the interval $(-\pi,\pi]$. (Harmonics $e^{it\omega_j}$ with frequencies $2\pi j/n$ outside this interval cannot be distinguished on the basis of observations at integer times only.) The frequencies $\omega_j=2\pi j/n$, $(-\pi,\pi]$. are called the Fourier frequencies of the series $\{x_1,\dots x_n\}$. The representation (3.1) can be rewritten in vector form as

$$\mathbf{x} = \sum_{j\in F_n} a_j e_j,$$ where

$$\mathbf{e}_j = n^{-1/2}(e^{i\omega_j}, e^{i2\omega_j}, e^{i3\omega_j}, \dots, e^{in\omega_j})', j \in F_n,$$
$$F_n = \{j \in \mathbb{Z} : -\pi < \omega_j \equiv 2\pi j/n \le \pi\} = \{-[(n-1)/2], \dots [n/2]\},$$

and [x] denotes the integer part of x. Notice that F_n contains n integers

Definition 3.3

(*The periodogram* of $\mathbf{x} \in \mathbb{C}^n$) The value $\mathbf{J}(\omega_j)$ of the periodogram of $\mathbf{x}$, at frequency $\omega_j=2\pi j/n$, $j \in F_n$, is defined in terms of the discrete Fourier transform $\{\alpha_j\}$ of $\mathbf{x}$ by,

$$\mathbf{J}(\omega_j) := |a_j|^2 = |\langle \mathbf{x}, \mathbf{e}_j \rangle|^2 = n^{-1}\left|\sum_{t=1}^{n} x_t e^{-it\omega_j}\right|^2.$$

Note that the periodogram decomposes $\|\mathbf{x}\|^2$ into a sum of components $|\langle \mathbf{x}, \mathbf{e}_j \rangle|^2$ associated with the Fourier frequencies, ω_j, $j \in F_n$. Thus

$$\|\mathbf{x}\|^2 = \sum_{j\in F_n} \mathbf{J}(\omega_j).$$

Tthe extension to the bivariate and multivariate case can be formulated proportionately. For the set of $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\} \in \mathbb{R}^2$ multivariate time series, the discrete Fourier transform is defined by

$$\mathbf{J}(\omega_j) = n^{-1/2}\, \mathbf{J}(\omega_j) = \begin{pmatrix} J_1 \\ J_2 \end{pmatrix} = n^{-1/2} \sum_{t=1}^{n} \mathbf{X}_t e^{-it\omega_j} \quad ,$$

https://doi.org/10.26219/heal.aueb.5789

where $\mathbf{J}_k(\omega_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_{kt} e^{-it\omega_j}$, and $\omega_j=2\pi j/n$, $-[(n-1)/2] \leq j \leq [n/2]$, are the n Fourier frequencies. The periodogram of $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ is defined at each of these frequencies ω_j to be the matrix

$$(\mathbf{J}(\omega))_{ij} = \mathbf{J}_i(\omega)\mathbf{J}_j^{\,*}(\omega)$$

where * denotes complex conjugate transpose. The definition is extended to all frequencies $\omega \in [-\pi, \pi]$ by setting

$$\mathbf{J}_{\mathrm{n}}(\omega) = \begin{cases} \mathbf{J}_{\mathrm{n}}(g(n,\omega)) & \text{if } \omega \geq 0 \\ \mathbf{J}_{\mathrm{n}}^{\,*}(g(n,-\omega)) & \text{if } \omega < 0, \end{cases}$$

where $g(n,\omega)$, $0 \leq \omega \leq \pi$, is the multiple of $2\pi/n$ closest to ω (the smallest one if two) or more simply

$$\mathbf{J}_n(\omega) = \begin{cases} \mathbf{J}_n(\omega_k), & \text{if } \omega_\kappa - \pi/n < \omega \leq \omega_\kappa + \pi/n \\ \mathbf{J}_n(-\omega), & \text{if } \omega \in [-\pi, 0] \end{cases}.$$

This definition shows that $\mathbf{J}_n$ is an even function that coincides with the original definition of the periodogram at the nonnegative Fourier frequencies.

We shall write $\mathbf{J}_{ij}(\omega)$, $i,j=1,\ldots,n$ for the components of $\mathbf{J}_n(\omega)$. Observe that $\mathbf{J}_{ii}(\omega)$ is the periodogram of the univariate observations $\{X_{1i}, \ldots, X_{ni}\}$. The function $\mathbf{J}_{12}(\omega)$ is called **the cross periodogram**. At the Fourier frequency ω_κ it has the value,

$$\mathbf{J}_{12}(\omega_\kappa) = n^{-1}\left(\sum_{t=1}^{n} X_{t1} e^{-it\omega_\kappa}\right)\left(\sum_{t=1}^{n} X_{t2} e^{it\omega_\kappa}\right).$$

Asymptotic properties of $\mathbf{J}_n$ can be deduced and they can be found extensively in Brockwell and Davis (1991).

Theorem 3.2

Let $\{\mathbf{X}_t\}$ be the linear process defined by $\mathbf{X}_t = \sum_{k=-\infty}^{\infty} C_k \mathbf{Z}_{t-k}$, $\{\mathbf{Z}_t\} \sim IID(\mathbf{0}, \boldsymbol{\Sigma})$, $\sum_{\kappa=-\infty}^{\infty} |\mathbf{C}_k(i,j)| |k|^{1/2} < \infty, i,j = 1,2$,where Σ is non singular, with periodogram $\mathbf{J}_n(\lambda) = \left[J_{ij}(\lambda)\right]_{i,j=1}^{2}, -\pi \leq \lambda \leq \pi.$

 https://doi.org/10.26219/heal.aueb.5789

i. If $0<\lambda_1<\lambda_2...<\lambda_m<\pi$ then the matrices $\mathbf{J}_n(\lambda_1),.....,\mathbf{J}_n(\lambda_m)$ converge jointly in distribution as $n\to\infty$ to independent random matrices, the k$^{\text{th}}$ of which is distributed as $\mathbf{W_k W_\kappa^*}$ where $W_k \sim N(\mathbf{0},2\pi f(\lambda_\kappa))$ and f is the spectral density matrix of $\{\mathbf{X_t}\}$.

ii. If $\omega_j=2\pi j/n\in[0,\pi]$ and $\omega_k=2\pi k/n\in[0,\pi]$

$$Cov\left(I_{pq}(\omega_j),I_{rs}(\omega_k)\right)=\begin{cases}(2\pi)^2\left[f_{pr}(\omega_j)f_{sq}(\omega_j)+f_{ps}(\omega_j)f_{qr}(\omega_j)\right]+O\left(n^{-1/2}\right) & if\ \omega_j=\omega_k=0\ or\ \pi\\ (2\pi)^2 f_{pr}(\omega_j)f_{sq}(\omega_j)+O\left(n^{-1/2}\right) & if\ 0<\omega_j=\omega_k<\pi\\ O\left(n^{-1}\right) & if\ \omega_j\neq\omega_\kappa\end{cases}$$

where the terms $O\left(n^{-1/2}\right)$ and $O\left(n^{-1}\right)$ can be bounded uniformly in j and k by $c_1n^{-1/2}$ and c_2n^{-1} respectively for some positive constants c_1,c_2.

3.3 Smoothing the periodogram

A good estimator $\hat{\theta}_n$ of any parameter θ should at least be consistent, i.e. should converge in probability to θ as $n\to\infty$. However $I_n(\lambda)/2\pi$ is not a consistent estimator of $f(\lambda)$. Since for large n the periodogram ordinates are approximately uncorrelated with variances changing only slightly over small frequency intervals, we might hope to construct a consistent estimator of $f(\lambda)$ by averaging the periodogram ordinates in a small neighborhood of λ. The number of Fourier frequencies in a given interval increases approximately linearly with *n*. By averaging the periodogram ordinates over a suitably increasing number of frequencies in a neighbourhood of λ, we can indeed construct consistent spectral density estimators as shown in the this section.

If $\mathbf{J}_n(\omega_j), j\in F_{n,}$ is the periodogram based on $\{\mathbf{X_1},\dots,\mathbf{X_n}\}$ then we may write

$$(2\pi)^{-1}\mathbf{J}_n(\omega_j)=f(\omega_j)+\mathbf{U}_jf(\omega_j),\quad j=1,...[(n-1)/2],$$

where f is defined by $f(\omega)=(2\pi)^{-1}\sum_{k=-\infty}^{\infty}\Gamma(k)e^{-ik\omega}$, $\omega\in[-\pi,\pi]$ and the sequence $\{\mathbf{U}_j\}$ is approximately $WN(\mathbf{0},\mathbf{I})$ for large *n*. In other words, we may think of

https://doi.org/10.26219/heal.aueb.5789

$(2\pi)^{-1}\mathbf{J}(\omega_j)$, $j=1,...[(n-1)/2]$ as an uncorrelated time series with a trend $f(\omega_j)$ which we wish to estimate.

A consistent estimator of the spectral matrix of a linear process with the form of {X_t}, can be obtained by smoothing the periodogram. Let {m} and {$W_n(\cdot)$} be sequences of integers and (scalar) weight functions respectively, satisfying conditions

$$\begin{aligned} & m \to \infty, \textit{and } m/n \to 0, \\ & W_n(k) = W_n(-k), W_n(k) \geq 0, \textit{for all } k \\ & \sum_{|k| \leq m} W_n(k) = 1, \\ & \sum_{|k| \leq m} W_n^{\ 2}(k) \to 0, \textit{as } n \to \infty \end{aligned} \tag{3.4}$$

We define the discrete spectral average estimator $\hat{f}$ by

$$\hat{f}(\omega) := (2\pi)^{-1} \sum_{|k| \leq m_n} W_n \mathbf{J}_n(g(n,\omega) + \omega_\kappa), 0 \leq \omega \leq \pi. \tag{3.5}$$

In order to evaluate $\hat{f}$ (ω), $0 \leq \omega \leq \pi$, we define $\mathbf{J}_n$ to have period 2π and replace $\mathbf{J}_n(0)$ whenever it appears in (3.5) by

$$\hat{f}(0) := (2\pi)^{-1} \operatorname{Re}\{W_n(0)\mathbf{J}_n(\omega_1) + 2\sum_{k=1}^{m} W_n(k)\mathbf{J}_n(\omega_{k+1})\}$$

We have applied the same weight function to all four components of $\mathbf{J}_n(\omega)$ in order to facilitate the statement and derivation of the properties of $\hat{f}$ (ω). It is frequently advantageous however to choose a different weight-function sequence for each component of $\mathbf{J}_n(\cdot)$ since the components may have quite diverse characteristics. Priestley (1981) provides a discussion of choosing weight functions to match the characteristics of $\mathbf{J}_n(\cdot)$.

The following theorem assets the consistency of the estimator $\hat{f}$ (ω).

Theorem 3.3

If {$\mathbf{X_t}$} is the linear process, $\mathbf{X}_t = \sum_{k=-\infty}^{\infty} C_k \mathbf{Z}_{t-k}$, $\{\mathbf{Z}_t\} \sim IID(\mathbf{0}, \mathbf{\Sigma})$, $\sum_{\kappa=-\infty}^{\infty} |\mathbf{C}_k(i,j)| |k|^{1/2} < \infty, i,j = 1,2$, where $\mathbf{\Sigma}$ is non singular and $\hat{f}(\omega) = [\hat{f}_{ij}(\omega)]^2_{i,j=1}$ is

https://doi.org/10.26219/heal.aueb.5789

the discrete spectral average estimator defined by $\hat{f}(\omega) := (2\pi)^{-1} \sum_{|k| \le m_n} W_n \mathbf{J}_n(g(n,\omega)+\omega_\kappa), 0 \le \omega \le \pi.$ then for $\lambda,\omega \in [0,\pi]$,

$$(\alpha) \lim_{n\to\infty} E\hat{f}(\omega) = f(\omega)$$

and

$$(b) \lim_{n\to\infty} \frac{Cov(\hat{f}_{pq}(\omega), f_{rs}(\lambda))}{\sum_{|k|\le m} W_n^{\,2}(k)} = \begin{cases} f_{pr}(\omega) f_{sq}(\omega) + f_{ps}(\omega) f_{qr}(\omega) & \text{if } \omega = \lambda = 0 \text{ or } \pi \\ f_{pr}(\omega) f_{sq}(\omega) & \text{if } 0 < \omega = \lambda < \pi \\ 0 & \text{if } \omega \ne \lambda \end{cases}$$

In choosing the $\{W_n(\cdot)\}$ (scalar) weight functions it is necessary to compromise between bias and variance of the spectral estimator. A weight function which assigns roughly equal weights to a broad band of frequencies will produce an estimate of $f(\cdot)$ which although smooth, may have a large bias, since the estimate of $f(\omega)$ depends on values of I_n at frequencies distant from ω. On the other hand a weight function which assigns most of its weight to a narrow frequency band centered at zero will give an estimator with relatively small bias, but with a large variance. In practice it is advisable to experiment with a range of weight functions and to select the one which appears to strike a satisfactory balance between bias and variance.

In the univariate case,the spectral density *f* is often estimated by a function of the form,

$$\hat{f}_L(\omega) = (2\pi)^{-1} \sum_{|h|\le r} w(h/r)\hat{\gamma}(h) e^{-ih\omega} , \qquad (3.6)$$

where $\gamma(\cdot)$ is the sample autocovariance function and $w(x)$ is an even, piecewise continuous function of x satisfying the conditions,

$$w(0) = 1,$$
$$|w(x)| \le 1, \quad for \quad all \quad x,$$
$$w(x) = 0, \quad for \quad |x| > 1$$

The function $w(\cdot)$ is called the lag window, and the corresponding estimator $\hat{f}_L$ is called the lag window spectral density estimator. If we assume that $r = r_n$ is a function of n such that $r \to \infty$ and $r/n \to 0$ as $n \to \infty$ then $\hat{f}_L$ is a sum of (2r+1) terms, each with a variance which is $O(n^{-1})$. If $\{r_n\}$ satisfies these conditions and

https://doi.org/10.26219/heal.aueb.5789

$\{X_t\}$ satisfies the conditions of the Theorem 3.3 then it can be shown that $\hat{f}_L$ is in fact a mean square consistent estimator of $f(\omega)$.

Although the discrete spectral average estimator and the lag window estimator have different expressions, they can be linked by means of the so called ***spectral window*** defined as,

$$W(\omega) = (2\pi)^{-1} \sum_{|h| \le r} w(h/r) e^{-ih\omega} , \tag{3.7}$$

We shall now present some of the examples of spectral windows that *Brockwell & Davis* (1991) provide.

Rectangular or Truncated Window

This window has the form

$$w(x) = \begin{cases} 1, & if \quad |x| \le 1 \\ 0 & otherwise, \end{cases}$$

and the corresponding spectral window is given by the Dirichlet kernel,

$$W(\omega) = (2\pi)^{-1} \frac{\sin((r+1/2)\omega)}{\sin(\omega/2)} .$$

Observe that $W(\omega)$ is negative for certain values of ω. This may lead to negative estimates of spectral density at certain frequencies. Then we have, as $n \to \infty$,

$$Var(\hat{f}_L(\omega)) \approx \frac{2r}{n} f^2(\omega) \quad for \quad 0 < \omega < \pi .$$

The Bartlett (1955) or Triangular Window

In this case

$$w(x) = \begin{cases} 1 - |x|, & if \quad |x| \le 1 \\ 0 & otherwise, \end{cases}$$

and the corresponding spectral window is given by the Fejer kernel,

$$W(\omega) = (2\pi r)^{-1} \frac{\sin^2(r\omega/2)}{\sin^2(\omega/2)} .$$

Since $W(\omega) \ge 0$, this window always gives non-negative spectral density estimates. Moreover, as $n \to \infty$,

https://doi.org/10.26219/heal.aueb.5789

$$Var(\hat{f}_L(\omega)) \sim \frac{r}{n} f^2(\omega)\int_{-1}^{1} w^2(x)d(x) = \frac{2r}{3n} f^2(\omega), \quad 0 < \omega < \pi .$$

The asymptotic variance is thus smaller than that of the rectangular lag window estimator using the same sequence $\{r_n\}$.

The Daniel Window

In this case the spectral window

$$W(\omega) = \begin{cases} \rho / 2\pi, & |\omega| \le \pi / r \\ 0, & otherwise, \end{cases}$$

corresponds to the discrete average estimator with weights

$$W_n(j) = (2m+1)^{-1}, \quad |j| \le m = [r/2n].$$

From (2.6) we find that the lag window corresponding to $W(\omega)$ is

$$w(h/r) = \int_{-\pi}^{\pi} W(\omega)e^{ih\omega}d\omega = \pi^{-1}(r/h)\sin(\pi h/r),$$

i.e.

$$w(x) = \sin(x\pi)/(x\pi), \quad -1 \le x \le 1 .$$

The corresponding lag window estimator has the asymptotic variance

$$Var(\hat{f}_L(\omega)) \sim rf^2(\omega)/n, \quad 0 < \omega < \pi .$$

Other windows are the Blackman-Tukey Window and the Parzen Window

3.4 Comparison of Lag-Window Estimators

Lag-Window Estimators may be compared by examining the spectral windows when the values of r for the different estimators are chosen in such a way that the estimators have the same asymptotic variance. Thus to compare the Bartlett and Daniell estimators we plot the spectral windows

$$W_B(\omega) = (2\pi r)^{-1}\sin^2(r\omega/2)/(\sin^2(\omega/2)) \text{ and } W_D(\omega) = r'/(2\pi), \quad |\omega| \le \pi / r'$$

where $r' = 2r/3$. The mass of the window W_B is spread over a broader frequency interval and has secondary peaks or “side-lobes” at some distance from the center. This means that Bartlett estimator with the same asymptotic variance as the Daniell

https://doi.org/10.26219/heal.aueb.5789

estimator is liable (depending on the spectral density being estimated) to exhibit greater bias

The width of the rectangular spectral window which leads to the same asymptotic variance as a given lag-window estimator is sometimes called the bandwidth of the given estimator. For example the Bartlett estimator with parameter r has bandwidth $2\pi / r' = 3\pi / r$.

3.5 Confidence Intervals for the Spectrum

In this section we provide two approximations to the distributions of the discrete spectral average estimator $\hat{f}(\omega)$ from which confidence intervals for the spectral density $f(\omega)$ can be constructed. Assume that $\{\mathbf{X}_t\}$ satisfies the conditions:

$$\mathbf{X}_t = \sum_{j=-\infty}^{j=\infty} \psi_j \mathbf{Z}_{t-j}, \quad \sum |\psi_j| |j|^{1/2} < \infty$$
$$\{\mathbf{Z}_t\} \sim IID(\mathbf{0}, \sigma^2)$$
$$E\mathbf{Z}^4_t < \infty$$

and that $\hat{f}$ is the discrete spectral average

$$\hat{f}(\omega_j) = (2\pi)^{-1} \sum_{|k| \le m} W_n(k) \mathbf{J}_n(\omega_j + \omega_k).$$

The χ^2 Aproximation

The random variables $\mathbf{I}_n(\omega_j + \omega_k) / (\pi f(\omega_j + \omega_k)), \quad -j < k < n/2 - j$

are approximately independent and distributed as chi-squared with 2 degrees of freedom. This suggests approximating the distribution of $\hat{f}(\omega_j)$ by the distribution of the corresponding linear combination of independent and identically distributed $\chi^2(2)$ random variables. However, as advocated by Tukey (1949), this distribution may in turn be approximated by the distribution of cY, where c is a constant, $Y \sim \chi^2(v)$ and c and v are found by the method of moments, i.e. by setting the mean and variance of cY, equal to the asymptotic mean and variance of $\hat{f}(\omega_j)$. This procedure gives the equations

https://doi.org/10.26219/heal.aueb.5789

$$cv = f(\omega_j)$$
$$2c^2 v = \sum_{|k|\le m} W^2{}_n(k) f^2(\omega_j)$$

from which we have that $c = \sum_{|k|\le m} W^2{}_n(k) f(\omega_j)/2$ and $v = 2/\sum_{|k|\le m} W^2{}_n(k)$. The number ν is called the ***equivalent degrees of freedom*** of the estimator $\hat{f}$. The distribution of $v\hat{f}(\omega_j)/f(\omega_j)$ is thus approximated by the chi-squared distribution with ν degrees of freedom, and interval

$$\left(\frac{v\hat{f}(\omega_j)}{\chi^2{}_{.975}(v)}, \frac{v\hat{f}(\omega_j)}{\chi^2{}_{.025}(v)}\right), \quad 0 < \omega_j < \pi, \tag{3.8}$$

is an approximate 95% confidence interval for $f(\omega_j)$. By taking logarithms in (3.8) we obtain the 95% confidence interval

$$(\ln \hat{f}(\omega_j) + \ln v - \ln \chi^2{}_{.975}(v), \ln \hat{f}(\omega_j) + \ln v - \ln \chi^2{}_{.025}(v)), \quad 0 < \omega_j < \pi, \qquad \text{for}$$

$\ln \hat{f}(\omega_j)$. This interval, unlike (3.8) has the same width for each $\omega_j \in (0,\pi)$.

The Normal Approximation

There are two intuitive justifications for making a normal approximation to the distribution of $\hat{f}(\omega_j)$. The first is that if the equivalent number of degrees of freedom v is large (i.e. if $\sum_{|k|\le m} W^2{}_n(k)$ is small) and if Y can be well approximated by a normal distribution with mean $cv = f(\omega_j)$ and variance $2c^2 v = \sum_{|k|\le m} W^2{}_n(k) f^2(\omega_j), \quad 0 < \omega_j < \pi$. The second is that we may approximate $\hat{f}(\omega_j)$ for n large by a sum of (2m+1) independent random variables, which by Lindeberg condition, is $\mathbf{AN}(f(\omega_j), \sum_{|k|\le m} W^2{}_n(k) f^2(\omega_j))$. Both points of view lead to the approximation for the distribution of $\hat{f}(\omega_j)$. Using this approximation we obtain the approximate 95% confidence bounds,

$$\hat{f}(\omega_j) \pm 1.96 \left(\sum_{|k|\le m} W^2{}_n(k)\right)^{1/2} \hat{f}(\omega_j) \text{ for } f(\omega_j).$$

https://doi.org/10.26219/heal.aueb.5789

Since the width of the confidence interval depends on $\hat{f}(\omega_j)$, it is customary to construct a confidence interval for $\ln f(\omega_j)$. The normal approximation to $\hat{f}(\omega_j)$ implies that $\ln \hat{f}(\omega_j)$ is $\mathbf{AN}(\ln f(\omega_j), \sum_{|k|\le m} W^2{}_n(k))$. Approximate 95% confidence bounds for $\ln f(\omega_j)$ are therefore given by

$$\ln \hat{f}(\omega_j) \pm 1.96\left(\sum_{|k|\le m} W^2{}_n(k)\right)^{1/2}.$$

3.5 Estimation of the Cross-Amplitude Spectrum

To estimate $a_{12}(\omega) = |f_{12}(\omega)| = |c_{12}(\omega) - iq_{12}(\omega)|$ we shall use $\hat{a}_{12}(\omega) := \left(\hat{c}_{12}{}^2(\omega) + \hat{q}_{12}{}^2(\omega)\right)^{1/2} = h\left(\hat{c}_{12}(\omega), \hat{q}_{12}(\omega)\right)$.

By the above conclusion of the asymptotic normal nature of $c_{12}(\omega)$, $q_{12}(\omega)$ and by $h(x, y) = \left(x^2 + y^2\right)^{1/2}$, we define that for $a_{12}(\omega)$>0, then

$$\hat{a}_{12}(\omega) \sim AN(a_{12}(\omega), a_n{}^2\sigma_a{}^2(\omega)),$$

where $a_n{}^2 = \sum_{|k|\le m} W^2(k)$ and

$$\sigma_a{}^2(\omega) = \left(\frac{\partial h}{\partial x}\right)^2 v_{33} + \left(\frac{\partial h}{\partial y}\right)^2 v_{44} + \left(\frac{\partial h}{\partial x}\right)\left(\frac{\partial h}{\partial y}\right) v_{34},$$

v_{ij} is the (i,j)-element of the matrix defined by (3.9) below and the derivatives of h are evaluated at $(c_{12}(\omega), q_{12}(\omega))$. Calculating the derivatives and simplifying, we find that if the squared coherency, $|K_{12}(\omega)|^2$, is strictly positive then

$$\hat{a}_{12}(\omega) \sim AN(a_{12}(\omega), a_n{}^2 a^2{}_{12}(\omega)\left(|K_{12}(\omega)|^{-2} + 1\right)/2).$$

Observe that for small values of $|K_{12}(\omega)|^2$, the asymptotic variance of $\hat{a}_{12}(\omega)$ is large.

The ***cospectrum*** $c_{12}(\omega) = [f_{12}(\omega) + f_{21}(\omega)]/2$ and the ***quadrature spectrum*** $q_{12}(\omega) = i[f_{12}(\omega) - f_{21}(\omega)]/2$ will be estimated by

https://doi.org/10.26219/heal.aueb.5789

$$\hat{c}_{12}(\omega) = [\hat{f}_{12}(\omega) + \hat{f}_{21}(\omega)]/2$$

and

$$\hat{q}_{12} = i[\hat{f}_{12}(\omega) - \hat{f}_{21}(\omega)]/2$$

respectively. The real valued random vector

$$(\sum_{|k|\le m} W_n{}^2(k))^{-1}\left(\hat{f}_{11}(\omega), \hat{f}_{22}(\omega), \hat{c}_{12}(\omega), \hat{q}_{12}(\omega)\right)', \; 0 < \omega < \pi \;,$$

has the asymptotic covariance matrix,

$$\mathbf{V} = \begin{bmatrix} f_{11}{}^2 & |f_{12}|^2 & f_{11}c_{12} & f_{11}q_{12} \\ |f_{12}|^2 & f_{22}{}^2 & f_{22}c_{12} & f_{22}q_{12} \\ f_{11}c_{12} & f_{22}c_{12} & \frac{1}{2}(f_{11}f_{22} + c_{12}{}^2 - q_{12}{}^2) & c_{12}q_{12} \\ f_{11}q_{12} & f_{22}q_{12} & c_{12}q_{12} & \frac{1}{2}(f_{11}f_{22} + q_{12}{}^2 - c_{12}{}^2) \end{bmatrix} \quad (3.9),$$

where the argument ω has been suppressed. Moreover we can express $\left(\hat{f}_{11}(\omega), \hat{f}_{22}(\omega), \hat{c}_{12}(\omega), \hat{q}_{12}(\omega)\right)'$ as the sum of (2m+1) random vectors,

$$\begin{bmatrix} \hat{f}_{11}(\omega) \\ \hat{f}_{22}(\omega) \\ \hat{c}_{12}(\omega) \\ \hat{q}_{12}(\omega) \end{bmatrix} = \sum_{|k|\le m} \mathbf{W}_n(k) \begin{bmatrix} \mathbf{I}_{11}(g(n,\omega)+\omega_k) \\ \mathbf{I}_{22}(g(n,\omega)+\omega_k) \\ \operatorname{Re}\{\mathbf{I}_{12}(g(n,\omega)+\omega_k) \\ -\operatorname{Im}\{\mathbf{I}(g(n,\omega)+\omega_k)\} \end{bmatrix},$$

where the summands are asymptotically independent. This suggests that

$$\left(\hat{f}_{11}(\omega), \hat{f}_{22}(\omega), \hat{c}_{12}(\omega), \hat{q}_{12}(\omega)\right)' \; is \; AN\left((f_{11}(\omega), f_{22}(\omega), c_{12}(\omega), q_{12}(\omega))', a_n{}^2\mathbf{V}\right) \quad (3.10)$$

where $a_n{}^2 = \sum_{|k|\le m} W_n{}^2(k)$ and $\mathbf{V}$ is defined by (3.9).

3.6 Estimation of the phase spectrum

The phase spectrum $\phi_{12}(\omega) = \arg f_{12}(\omega)$ will be estimated by

$$\hat{\phi}_{12}(\omega) := \arg(\hat{c}_{12}(\omega) - i\hat{q}_{12}(\omega)) \in (-\pi, \pi].$$

If $|K_{12}(\omega)|^2$>0, then by (3.10),

https://doi.org/10.26219/heal.aueb.5789

$$\hat{\phi}_{12}(\omega)\, is\ AN(\phi_{12}(\omega), a_n^{\ 2} a^2_{\ 12}(\omega)\left(\left|K_{12}(\omega)\right|^{-2} - 1\right)/2).$$

The asymptotic variance of $\hat{\phi}_{12}(\omega)$, like that of $\hat{a}_{12}(\omega)$, is large if $|K_{12}(\omega)|^2$ is small. In the case when $|K_{12}(\omega)|^2$ =0, both $c_{12}(\omega)$ and $q_{12}(\omega)$ are zero, so from (3.9), (3.10)

$$\begin{bmatrix} \hat{c}_{12}(\omega) \\ \hat{q}_{12}(\omega) \end{bmatrix} is\ AN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \frac{1}{2} a^2{}_n \begin{bmatrix} f_{11} f_{22} & 0 \\ 0 & f_{11} f_{22} \end{bmatrix}\right).$$

As $\hat{\phi}_{12}(\omega) = \arg(\hat{c}_{12}(\omega) - i\hat{q}_{12}(\omega)) = \arg\left[\left(a_n^{\ 2} f_{11} f_{22}/2\right)^{-1/2}\left(\hat{c}_{12}(\omega) - i\hat{q}_{12}(\omega)\right)\right]$, we conclude that $\hat{\phi}_{12}(\omega) \Rightarrow \arg\left[U_1 + iU_2\right]$, where U_1, U_2 are independent standard normal variables, it can be shown that $\arg\left[U_1 + iU_2\right]$ is uniformly distributed on (-π,π). Hence if n is large and $|K_{12}(\omega)|$=0, $\hat{\phi}_{12}(\omega)$ is approximately uniformly distributed on (-π,π).

Since $\hat{\phi}_{12}(\omega) \sim AN(\phi_{12}(\omega), a_n^{\ 2} a^2_{\ 12}(\omega)\left(\left|K_{12}(\omega)\right|^{-2} - 1\right)/2)$ we obtain the approximate 95% confidence bounds for $\phi_{12}(\omega)$,

$$\hat{\phi}_{12}(\omega) \pm 1.96 a_n \hat{a}_{12}(\omega)\left(\left|\hat{K}_{12}(\omega)\right|^{-2} - 1\right)^{1/2} / 2^{1/2},$$

where $\left|\hat{K}_{12}(\omega)\right|^2$ is the estimated squared coherency,

$$\left|\hat{K}_{12}(\omega)\right|^2 = \hat{a}_{12}^2(\omega) / \left[\hat{f}_{11}(\omega)\hat{f}_{22}(\omega)\right],$$

and it is assumed that $\left|\hat{K}_{12}(\omega)\right|^2$>0.

Hannan (1970), p.257 discusses an alternative method for constructing a confidence region for $\phi_{12}(\omega)$ in the case when $W_n(k) = (2m+1)^{-1}$ for $|k| \le m$ and $W_n(k) = 0$ for $|k| > m$. He shows that if the distribution of the periodogram is replaced by the asymptotic distributions of Theorem 3.1, then the event E has probability $(1-a)$, where

https://doi.org/10.26219/heal.aueb.5789

$$E = \left\{ \left| \sin\left(\hat{\phi}_{12}(\omega) - \phi_{12}(\omega)\right) \right| \leq \left[\frac{1 - \left|\hat{K}_{12}(\omega)\right|^2}{4m\left|\hat{K}_{12}(\omega)\right|^2} \right]^{1/2} t_{1-a/2}(4m) \right\}$$

and $t_{1-a/2}(4m)$ is the $(1-a/2)$-quantile of the t-distribution with $4m$ degrees of freedom. For given values of $\hat{\phi}_{12}(\omega)$ and $\left|\hat{K}_{12}(\omega)\right|$, the set of $\phi_{12}(\omega)$ values satisfying the inequality which defines E is therefore a $100(1-a)\%$ confidence region for $\phi_{12}(\omega)$. If the right hand side of the inequality is greater than or equal to 1 (as will be the case if $\left|\hat{K}_{12}(\omega)\right|$ is sufficient small), then we obtain the uninformative confidence interval $(-\pi, \pi]$ for $\phi_{12}(\omega)$. On the other hand if the right-hand side is less than one, let us denote its arcsin (in $[0, \pi/2)$) by ϕ^*. Our confidence region then consists of values $\phi_{12}(\omega)$ such that

$$\left| \sin\left(\hat{\phi}_{12}(\omega) - \phi_{12}(\omega)\right) \right| \leq \sin \phi^*$$

i.e. such that

$$\hat{\phi}_{12}(\omega) - \phi^* \leq \phi_{12} \leq \hat{\phi}_{12}(\omega) - \phi^*$$

or

$$\hat{\phi}_{12}(\omega) + \pi - \phi^* \leq \phi_{12} \leq \hat{\phi}_{12}(\omega) + \pi - \phi^*.$$

The confidence region can thus be represented as a union of two subintervals of the unit circle whose centers are diametrically opposed (at $\hat{\phi}_{12}(\omega)$ and $\hat{\phi}_{12}(\omega) + \pi$) and whose arc lengths are $2\phi^*$. If $\left|\hat{K}_{12}(\omega)\right|^2$ is close to one, then we normally choose the interval centered at $\hat{\phi}_{12}(\omega)$, since the other interval corresponds to a sign change in both $c_{12}(\omega), q_{12}(\omega)$ which is unlikely if $\left|\hat{K}_{12}(\omega)\right|^2$ is close to one.

3.7 Estimation of the Absolute Coherence

The squared coherency $\left|K_{12}(\omega)\right|^2$ is estimated by $\left|\hat{K}_{12}(\omega)\right|^2$ where

$$\left|\hat{K}_{12}(\omega)\right| = \left[\hat{c}_{12}^2(\omega) + \hat{q}_{12}^2(\omega)\right]^{1/2} / \left[\hat{f}_{11}(\omega)\hat{f}_{22}(\omega)\right]^{1/2} = h(\hat{f}_{11}(\omega), \hat{f}_{22}(\omega), \hat{c}_{12}(\omega), \hat{q}_{12}(\omega))$$

If $\left|\hat{K}_{12}(\omega)\right|>0$ then

$$\left|\hat{K}_{12}(\omega)\right| \sim AN(\left|K_{12}(\omega)\right|, a_n^{\;2}(1-\left|K_{12}(\omega)\right|^2)^2/2)$$

giving the approximate 95% confidence bounds,

$$\left|\hat{K}_{12}(\omega)\right| \pm 1.96 a_n (1-\left|\hat{K}_{12}(\omega)\right|^2)/\sqrt{2},$$

for $\left|\hat{K}_{12}(\omega)\right|$.

Since $d\left[\tanh^{-1}(x)\right]/dx = d\left[\frac{1}{2}\ln\left(\frac{1+x}{1-x}\right)\right]/dx = \left(1-x^2\right)^{-1}$ it follows that

$$\tanh^{-1}\left(\left|\hat{K}_{12}(\omega)\right|\right) \sim AN\left(\tanh^{-1}\left(\left|K_{12}(\omega)\right|\right), a_n^{\;2}/2\right).$$

From this we obtain the constant-width large sample $100\left(1-a\right)\%$ confidence interval,

$$\left(\tanh^{-1}\left(\left|\hat{K}_{12}\right|\right) - \Phi_{(1-a/2)} a_n/\sqrt{2}, \tanh^{-1}\left(\left|\hat{K}_{12}\right|\right) + \Phi_{(1-a/2)} a_n/\sqrt{2}\right),$$

for $\tanh^{-1}\left(\left|\hat{K}_{12}\right|\right)$. The corresponding $100\left(1-a\right)\%$ confidence region for $\left|K_{12}(\omega)\right|$ is the intersection with [0,1] of the interval

$$\left(\tanh\left[\tanh^{-1}\left(\left|\hat{K}_{12}(\omega)\right|\right) + \Phi_{1-\alpha/2}\alpha_n/\sqrt{2}\right]\right),$$

assuming still that $\left|K_{12}(\omega)\right|>0$.

If the weight function $W_n(k)$ in (1) has the form $W_n(k) = \begin{cases} (2m+1)^{-1} & if\,|k| \le m \\ 0 & if\,|k| > m \end{cases}$,

then the hypothesis $\left|K_{12}(\omega)\right|=0$ can be tested against the alternative hypothesis $\left|K_{12}(\omega)\right|>0$ using the statistic,

$$Y = 2m\left|\hat{K}_{12}(\omega)\right|^2 / \left[1-\left|\hat{K}_{12}(\omega)\right|^2\right].$$

https://doi.org/10.26219/heal.aueb.5789

Under the approximating asymptotic distribution of Theorem 3.2, it can be shown that $\left|\hat{K}_{12}(\omega)\right|^2$ is distributed as the square of a multiple correlation coefficient, so that $Y \sim F(2,4m)$ under the hypothesis that $\left|K_{12}(\omega)\right| = 0$. We therefore reject the hypothesis $\left|K_{12}(\omega)\right| = 0$ if $Y \sim F_{1-a}(2,4m)$ is at east as large as the $(1-a)$-quantile of the F distribution with 2 and 4m degrees of freedom. The power of this test has been tabulated for numerous values of $\left|K_{12}(\omega)\right|$>0 by Amos and Koopmans (1963).

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

CHAPTER 4

COINTEGRATION

4.1 Spurious Regression and Cointergration

In the context of describing cointegration method and usage we have to mention *E. Zivot and Wang (2006)* and *K. Juselius (2001)*

4.1.1 Intergrated process

Def.4.1 *(Intergrated process)*

$\{y_t\}$ is an intergrated process of order 1, denoted $y_t \sim I(1)$, if it has the form

$$y_t = y_{t-1} + u_t \tag{4.1}$$

where u_t is a stationary time series. By definition, the first difference of y_t is stationary

$$\Delta y_t = u_t$$

Since the stationary process u_t does not need to be differenced, it is called an intergrated process of order zero and is denoted $u_t \sim I(0)$.

4.1.2 Cointegration

Let $\mathbf{Y}_t = (y_{1t}, ..., y_{nt})'$ denote a $(n \times 1)$ vector of $I(1)$ time series. $\mathbf{Y}_t$ is *cointegrated* if there exists an $(n \times 1)$ vector $\boldsymbol{\beta} = (\beta_1, ..., \beta_n)'$ such that

$$\boldsymbol{\beta}'\mathbf{Y} = \beta_1 y_{1t} + ... + \beta_n y_{nt} \sim I(0) \tag{4.2}$$

In other words, the nonstationary time series $\mathbf{Y}_t$ are cointegrated if there is a linear combination of them that is stationary or $I(0)$. The linear combination $\boldsymbol{\beta}'\mathbf{Y}_t$ is often motivated by economic theory and referred to as a long-run equilibrium relationship. The intuition is that $I(1)$ time series with a long-run equilibrium relationship cannot drift too far apart from the equilibrium because economic forces will act to restore the equilibrium relationship.

 https://doi.org/10.26219/heal.aueb.5789

Normalization

The cointegration vector β in (4.2) is not unique since for any scalar c the linear combination $c\boldsymbol{\beta}'\mathbf{Y}_t = \boldsymbol{\beta}^{*\prime}\mathbf{Y}_t \sim I(0)$. Hence, some *normalization* assumption is required to uniquely identify β. A typical normalization is

$$\boldsymbol{\beta}'\mathbf{Y}_t = y_{1t} - \beta_2 y_{2t} - .... - \beta_n y_{nt} \sim I(0)$$

or

$$y_{1t} = \beta_2 y_{2t} + ... + \beta_n y_{nt} + u_t \qquad (4.3)$$

where $u_t \sim I(0)$. In (4.3), the error term u_t is often referred to as the *disequilibrium error* or the *cointegrating residual*. In long-run equilibrium, the disequilibrium error u_t is zero and the long-run equilibrium relationship is

$$y_{1t} = \beta_2 y_{2t} + ... + \beta_n y_{nt} \qquad (4.4)$$

Multiple Cointegrating Relationships

If the $(n \times 1)$ vector $\mathbf{Y}_t$ is cointegrating there may be $0 < r < n$ linearly independent cointegrating vectors. For example, let $n = 3$ and suppose there are $r = 2$ cointegrating vectors $\beta_1 = (\beta_{11}, \beta_{12}, \beta_{13})'$ and $\beta_2 = (\beta_{21}, \beta_{22}, \beta_{23})'$. Then, $\beta_1'\mathbf{Y}_t = \beta_{11}y_{1t} + \beta_{12}y_{2t} + \beta_{13}y_{3t} \sim I(0)$ and $\beta_2'\mathbf{Y}_t = \beta_{21}y_{1t} + \beta_{22}y_{2t} + \beta_{23}y_{3t} \sim I(0)$ .The rows of the (3×2) matrix

$$\mathbf{B} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \end{pmatrix}$$

forms a *basis* for the space of cointegrating vectors. The linearly independent vectors β_1 and β_2 in cointegrating basis $\mathbf{B}$ are not unique unless some normalization assumptions are made. Furthermore, any linear combination of β_1 and β_2, e.g. $\beta_3 = c_1\beta_1 + c_2\beta_2$ where c_1 and c_2 are constants, is also a cointegrating vector.

4.1.3 Spurious Regression

The time series regression modelling requires all variables to be $I(0)$. In this case, the usual statistical results for the linear regression model hold. If some or all of

 https://doi.org/10.26219/heal.aueb.5789

the variables in the regression are $I(1)$ then the usual statistical results may not hold. One important case in which the usual statistical results do not hold is *spurious regression* when all regressions are $I(1)$ and not cointegrated.

Statistical Implications of Spurious Regression

Let $\mathbf{Y}_t = (y_{1t},....,y_{nt})'$ denote as $(n\times1)$ vector of $I(1)$ time series that are not cointegrated. Using the partition $\mathbf{Y}_t = (y_{1t},\mathbf{Y}'_{2t})'$, consider the least squares regression of y_{1t} on $\mathbf{Y}_{2t}$ giving the fitted model

$$y_{1t} = \hat{\beta}_2'\mathbf{Y}_{2t} + \hat{u}_t \tag{4.5}$$

Since y_{1t} is not cointegrated with $\mathbf{Y}_{2t}$, (3.1) is a *spurious regression* and there exists no true value of β_2 which means that the residuals are stationary. The following results about the behavior of $\hat{\beta}_2$ in the spurious regression (4.5) are due to Phillips (1986):

- $\hat{\beta}_2$ does not converge in probability to zero but instead converges in distribution to a non-normal random variable not necessarily centred at zero. This is the spurious regression phenomenon.
- The usual OLS *t*-statistics for testing that the elements of β_2 are zero diverge to $\pm\infty$ as $T\to\infty$. Hence, with large enough sample it will appear that $\mathbf{Y}_t$ is cointegrated, when it is not, if the usual asymptotic normal inference is used.
- The usual R^2 from the regression converges to unity as $T\to\infty$ so that the model will appear to fit well even though it is misspecified.

Regression with $I(1)$ data only makes sense when the data are cointegrated.

4.1.4 Cointegration and Common Trends

If the $(n\times1)$ vector time series $\mathbf{Y}_t$ is cointegrated with $0<r<n$ cointegrating vectors then there are $n-r$ *common* $I(1)$ *stochastic trends*. To illustrate the duality between cointegration and common trends, let $\mathbf{Y}_t = (y_{1t},y_{2t})' \sim I(1)$ and

https://doi.org/10.26219/heal.aueb.5789

$\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t})' \sim I(0)$ and suppose that $\mathbf{Y}_t$ is cointegrated with cointegrating vector $\boldsymbol{\beta} = (1, -\beta_2)'$. This cointegration relationship may be represented as

$$\begin{aligned} y_{1t} &= \beta_2 \sum_{s=1}^{t} \varepsilon_{1s} + \varepsilon_{3t} \\ y_{2t} &= \sum_{s=1}^{t} \varepsilon_{1s} + \varepsilon_{2t} \end{aligned} \tag{4.6}$$

The common stochastic trend is $\sum_{s=1}^{t} \varepsilon_{1s}$. Notice that the cointegrating relationship annihilates the common stochastic trend:

$$\boldsymbol{\beta}'\mathbf{Y}_t = \beta_2 \sum_{s=1}^{t} \varepsilon_{1s} + \varepsilon_{3t} - \beta_2 \left(\sum_{s=1}^{t} \varepsilon_{1s} + \varepsilon_{2t} \right) = \varepsilon_{3t} - \beta_2 \varepsilon_{2t} \sim I(0). \tag{4.7}$$

4.1.5 Simulating Cointegrating Systems

Cointegrating systems may be conveniently simulated using Phillips (1991) *triangular representation*. For example, consider a bivariate cointegrated system for $\mathbf{Y}_t = (y_{1t}, y_{2t})'$ with cointegrating vector $\boldsymbol{\beta} = (1, -\beta_2)'$. A triangular representation has the form

$$y_{1t} = \beta_2 y_{2t} + u_t, \text{ where } u_t \sim I(0) \tag{4.8}$$

$$y_{2t} = y_{2t-1} + v_t, \text{ where } v_t \sim I(0) \tag{4.9}$$

The first equation describes the long-run equilibrium relationship with an $I(0)$ disequilibrium error u_t. The second equation specifies y_{2t} as the common stochastic trend with innovation v_t:

$$y_{2t} = y_{20} + \sum_{j=1}^{t} v_j \tag{4.10}$$

In general, the innovations u_t and v_t may be contemporaneously and serially correlated. The time series structure of these innovations characterizes the short-run dynamics of the cointegrated system. The system (4.8)-(4.9) with $\beta_2 = 1$, for example, might be used to model the behaviour of the *logarithm* of spot and forward prices, spot and futures prices or stock prices and dividends.

https://doi.org/10.26219/heal.aueb.5789

4.2 Residual-Based Tests for Cointegration

Let the $(n \times 1)$ vector $\mathbf{Y}_t$ be $I(1)$. Recall, $\mathbf{Y}_t$ is cointegrated with $0 < r < n$ cointegrating vectors if there exists an $(r \times n)$ matrix $\mathbf{B}'$ such that

$$\mathbf{B}'\mathbf{Y}_t = \begin{pmatrix} \beta_1'\mathbf{Y}_t \\ \vdots \\ \beta_r'\mathbf{Y}_t \end{pmatrix} = \begin{pmatrix} u_{1t} \\ \vdots \\ u_{rt} \end{pmatrix} \sim I(0) \tag{4.11}$$

Testing for cointegration may be thought of as testing for the existence of long-run equilibria among the elements of $\mathbf{Y}_t$. Cointegration tests cover two situations:

- There is at most one cointegrating vector
- There are at most $0 \le r < n$ cointegrating vectors.

The first case is originally considered by Engle and Granger (1986) and they developed a simple two-step residual-based testing procedure based on regression techniques. The second case was originally considered by Johansen (1988) who developed a sophisticated sequential procedure for determining the existence of the cointegration and for determining the number of cointegrating relationships based on the maximum likelihood techniques. This section explains Engle and Granger's two-step procedure. Johansen's more general procedure will be discussed later on.

Engle and Granger's (1986) two-step procedure for determining if the $(n \times 1)$ vector β is a cointegrating vector is as follows:

- Form the cointegrating residual $\beta'\mathbf{Y}_t = u_t$
- Perform a unit root test on u_t to determine if it is $I(0)$.

The null hypothesis in the Engle and Granger procedure is no-cointegration and the alternative is cointegration. There are two cases to consider. In the first case, the proposed cointegrating vector β is pre-specified (not estimated). For example, economic theory may imply specific values for the elements in β such as $\beta = (1,-1)'$. The cointegrating residual is then readily constructed using the prespecified cointegrating vector. In the second case, it is proposed that the cointegrating vector is estimated from the data and an estimate of the cointegrating residual $\hat{\beta}'\mathbf{Y}_t = \hat{u}_t$ is formed. Tests for cointegration using a pre-specified cointegrating vector are generally much more powerful than the tests employing an estimated vector.

4.2.1 Testing for Cointegration When the Cointegrating Vector Is Pre-specified

Let $\mathbf{Y}_t$ denote an $(n \times 1)$ vector of $I(1)$ time series, let β be an $(n \times 1)$ prespecified cointegrating vector and let $u_t = \beta' \mathbf{Y}_t$ denote the prespecified cointegrating residual. The hypotheses to be tested are

$$H_0 : u_t = \beta' \mathbf{Y}_t \sim I(1) \text{ (no cointegration)} \tag{4.12}$$

$$H_1 : u_t = \beta' \mathbf{Y}_t \sim I(0)$$

Any unit root test statistic may be used to evaluate the above hypotheses. The most popular choices are the ADF (Dickey and Fuller (1979)) and PP statistics (Phillips-Perron (1988)). Cointegration is found if the unit root test rejects the no-cointegration null. It should be kept in mind, however, that the cointegrating residual may include deterministic terms (constant or trend) and the unit root tests should account for these terms accordingly.

4.2.2 Testing for Cointegration When the Cointegrating Vector is Estimated

Let $\mathbf{Y}_t$ denote an $(n \times 1)$ vector of $I(1)$ time series and let β denote an $(n \times 1)$ unknown cointegrating vector. The hypotheses to be tested are given in (4.12). Since β is unknown, to use the Engle-Granger procedure it must be first estimated from the data. Before β can be estimated some normalization assumption must be made to uniquely identify it. A common normalization is to specify the first element in $\mathbf{Y}_t$ as the dependent variable and the rest as the explanatory variables. Then $\mathbf{Y}_t = \left(y_{1t}, \mathbf{Y}_{2t}'\right)'$ where $\mathbf{Y}_{2t} = \left(y_{2t}, ..., y_{nt}\right)'$ is an $((n-1) \times 1)$ vector and the cointegrating vector is normalized as $\beta = \left(1, -\beta_2'\right)'$. Engle and Granger propose estimating the normalized cointegrating vector β_2 by least squares from the regression

$$y_{1t} = c + \beta_2' \mathbf{Y}_{2t} + u_t \tag{4.13}$$

and testing the no-cointegration hypothesis (4.12) with a unit root test using the estimated cointegrating residual

https://doi.org/10.26219/heal.aueb.5789

$$\hat{u}_t = y_{1t} - \hat{c} - \beta_2 \mathbf{Y}_{2t} \tag{4.14}$$

where $\hat{c}$ and $\hat{\beta}_2$ are the least squares estimates of c and β_2. The unit root test regression in this case is without deterministic terms (constant or constant and trend). Phillips and Ouliaris (1990) show that ADF and PP unit root tests (t-tests and normalized bias applied to the estimated cointegrating residual (4.14) do not have the usual Dickey-Fuller distributions under the null hypothesis (4.12) of no-cointegration. Instead, due to the spurious regression phenomenon under the null hypothesis (4.12), the distribution of the ADF and PP unit root tests have asymptotic distributions that are functions of Wiener processes that depend on deterministic terms in the regression (4.13) used to estimate β_2 and the number of variables, distributions, $n-1$, in $\mathbf{Y}_{2t}$. These distributions are known as the *Phillips-Ouliaris* (PO) distributions and are described in Phillips and Ouliaris (1990). To further complicate matters, Hansen (1992) showed the appropriate PO distributions of the ADF and PP unit root tests applied to the residuals (4.14) also depend on the trend behavior of y_{1t} and $\mathbf{Y}_{2t}$ as follows:

Case I: $\mathbf{Y}_{2t}$ and y_{1t} are both $I(1)$ without drift. The ADF and PP unit root test statistics follow the PO distributions, adjusted for a constant, with dimension parameter $n-1$.

Case II: $\mathbf{Y}_{2t}$ is $I(1)$ with drift and y_{1t} may not be $I(1)$ with drift. The ADF and PP unit root test statistics follow the PO distributions, adjusted for a constant and trend, with dimension parameter $n-2$. If $n-2=0$ then the ADF and PP unit root test statistics follow the DF distributions adjusted for a constant and trend.

Case III: $\mathbf{Y}_{2t}$ is $I(1)$ without drift and y_{1t} is $I(1)$ with drift. In this case, β_2 should be estimated from the regression

$$y_{1t} = c + \delta t + \beta_2' \mathbf{Y}_{2t} + u_t \tag{4.15}$$

The resulting ADF and PP unit root test statistics on the residuals from (4.15) follow the PO distributions, adjusted for a constant and trend, with dimension parameter $n-1$.

https://doi.org/10.26219/heal.aueb.5789

4.3 VAR Models and Cointegration

The Granger representation theorem links cointegration to error correction models. In a series of important papers and in a marvelous textbook, Soren Johansen firmly roots cointegration and error correction models in a vector autoregression framework. This section outlines Johansen's (1988) approach to cointegration modeling.

4.3.1 The Cointegrated VAR

Consider the levels of VAR(p) model for the $(n \times 1)$ vector $\mathbf{Y}_t$

$$\mathbf{Y}_t = \mathbf{\Phi D}_t + \mathbf{\Pi}_1 \mathbf{Y}_{t-1} + \cdots + \mathbf{\Pi}_p \mathbf{Y}_{t-p} + \varepsilon_t, \quad t = 1, \ldots, T, \tag{4.16}$$

where $\mathbf{D}_t$ contains deterministic terms (constant, trend, seasonal dummies etc.). Recall, the VAR(p) model is stable if

$$\det\left(\mathbf{I}_n - \mathbf{\Pi}_1 z - \cdots - \mathbf{\Pi}_p z^p\right) = 0 \tag{4.17}$$

has all roots outside the complex unit circle. If (4.17) has a root on the unit circle then some or all of the variable in $\mathbf{Y}_t$ are $I(1)$ and they may also be cointegrated. Recall, $\mathbf{Y}_t$ is cointegrated if there exists some linear combination of the variable in $\mathbf{Y}_t$ that is $I(0)$. Suppose $\mathbf{Y}_t$ is $I(1)$ and possibly cointegrated. Then the VAR representation (4.16) is not the most suitable representation for analysis because the cointegrating relations are not explicitly apparent. The cointegrating relations become apparent if the levels VAR (4.16) is transformed to the *vector error correction model* (VECM)

$$\Delta \mathbf{Y}_t = \mathbf{\Phi D}_t + \mathbf{\Pi Y}_{t-1} + \mathbf{\Gamma}_1 \Delta \mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{p-1} \Delta \mathbf{Y}_{t-p+1} + \varepsilon_t \tag{4.18}$$

where $\mathbf{\Pi} = \mathbf{\Pi}_1 + \cdots + \mathbf{\Pi}_p - \mathbf{I}_n$ and $\mathbf{\Gamma}_k = -\sum_{j=k+1}^{p} \mathbf{\Pi}_j$, $k = 1, \ldots, p-1$. The matrix $\mathbf{\Pi}$ is called the long-run impact matrix and $\mathbf{\Gamma}_k$ are the short-run impact matrices. Notice that the VAR parameters $\mathbf{\Pi}_i$ may be recovered from the VECM parameters $\mathbf{\Pi}$ and $\mathbf{\Gamma}_k$ via

$$\mathbf{\Pi}_1 = \mathbf{\Gamma}_1 + \mathbf{\Pi} + \mathbf{I}_n \tag{4.19}$$

$$\mathbf{\Pi}_k = \mathbf{\Gamma}_k - \mathbf{\Gamma}_{k-1}, \quad k = 2, \ldots, p.$$

https://doi.org/10.26219/heal.aueb.5789

In the VECM (4.18), $\Delta\mathbf{Y}_t$ and its lags are $I(0)$. The term $\mathbf{\Pi Y}_{t-1}$ is the only one which includes potential $I(1)$ variables and for $\Delta\mathbf{Y}_t$ to be $I(0)$ it must be the case that $\mathbf{\Pi Y}_{t-1}$ is also $I(0)$. Therefore, $\mathbf{\Pi Y}_{t-1}$ must contain the cointegrating relations if they exist. If the VAR(p) process has unit root then from (4.17) it is clear that $\mathbf{\Pi}$ is a singular matrix. If $\mathbf{\Pi}$ is singular then it has *reduced rank*; that is $rank(\mathbf{\Pi}) = r < n$. There are two cases to consider:

1. $rank(\mathbf{\Pi}) = 0$. This implies that $\mathbf{\Pi} = 0$ and $\mathbf{Y}_t$ is $I(1)$ and not cointegrated. The VECM (4.18) reduces to a VAR(p-1) in first differences

$$\Delta\mathbf{Y}_t = \mathbf{\Phi D}_t + \mathbf{\Pi Y}_{t-1} + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \varepsilon_t .$$

2. $0 < rank(\mathbf{\Pi}) = r < n$. This implies that $\mathbf{Y}_t$ is $I(1)$ with r linearly independent cointegrating vectors and $n-r$ common stochastic trends (unit roots). Since $\mathbf{\Pi}$ has rank r it can be written as the product

$$\underset{(n\times n)}{\mathbf{\Pi}} = \underset{(n\times r)}{\alpha}\,\underset{(r\times n)}{\beta}'$$

where α and β are $(n\times r)$ matrices with $rank(\alpha) = rank(\beta) = r$. The rows of β' form a basis for the r cointegrating vectors and the elements of α distribute the impact of the cointegrating vectors to the evolution of $\Delta\mathbf{Y}_t$. The VECM (4.18) becomes

$$\Delta\mathbf{Y}_t = \mathbf{\Phi D}_t + \alpha\beta'\mathbf{Y}_{t-1} + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \varepsilon_t , \tag{4.20}$$

where $\beta'\mathbf{Y}_{t-1} \sim I(0)$ since β' is a matrix of cointegrating vectors.

It is important to recognize that the factorization $\mathbf{\Pi} = \alpha\beta'$ is not unique since for any $r\times r$ nonsingular matrix $\mathbf{H}$ we have

$$\alpha\beta' = \alpha\mathbf{H}\mathbf{H}^{-1}\beta' = (\alpha\mathbf{H})\left(\beta\mathbf{H}^{-1\prime}\right)' = \alpha^*\beta^{*\prime}.$$

Hence the factorization $\mathbf{\Pi} = \alpha\beta'$ only identifies the space spanned by the cointegrating relations. To obtain unique values of α and β' requires further restrictions on the model.

https://doi.org/10.26219/heal.aueb.5789

4.3.2 Johansen's Methodology for Modeling Cointegration

The basic steps in Johansen's Methodology are:

- Specify and estimate a VAR(*p*) model for $\Delta\mathbf{Y}_t$.
- Construct likelihood ratio tests for the rank of $\mathbf{\Pi}$ to determine the number of cointegrating vectors.
- If necessary, impose normalization and identifying restrictions on the cointegrating vectors.
- Given the normalized cointegrating vectors estimate the resulting cointegrated VECM by maximum likelihood.

4.3.3 Specification of Deterministic Terms

Following Johansen (1995), the deterministic terms in (4.20) are restricted to the form

$$\mathbf{\Phi D}_t = \mu_t = \mu_0 + \mu_1 t$$

If the deterministic terms are unrestricted then the time series in $\mathbf{Y}_t$ may exhibit quadratic trends and there may be linear trend term in the cointegrating relationships. Restricted versions of the trend parameters μ_0 and μ_1 limit the trending nature of the series in $\mathbf{Y}_t$. The trend behavior of $\mathbf{Y}_t$ can be classified into five cases:

I. Model $H(r)$: $\mu_t = 0$ (no constant). The restricted VECM is

$$\Delta\mathbf{Y}_t = \alpha\beta'\mathbf{Y}_{t-1} + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \varepsilon_t, \qquad (4.21)$$

and all the series in $\mathbf{Y}_t$ are $I(1)$ without drift and the cointegrating relations $\beta'\mathbf{Y}_t$ have mean zero.

II. Model $H_1^*(r)$: $\mu_t = \mu_0 = \alpha\rho_0$ (ρ_0 is restricted constant). The restricted VECM is

$$\Delta\mathbf{Y}_t = \alpha\left(\beta'\mathbf{Y}_{t-1} + \rho_0\right) + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \varepsilon_t, \qquad (4.22)$$

the series in $\mathbf{Y}_t$ are $I(1)$ without drift and the cointegrating relations $\beta'\mathbf{Y}_t$ have non-zero means ρ_0.

III. Model $H_1(r)$: $\mu_t = \mu_0$ (unrestricted constant). The restricted VECM is

https://doi.org/10.26219/heal.aueb.5789

$$\Delta \mathbf{Y}_t = \mu_0 + \alpha\beta'\mathbf{Y}_{t-1} + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \varepsilon_t, \tag{4.23}$$

the series in $\mathbf{Y}_t$ are $I(1)$ with drift vector μ_0 and the cointegrating relations $\beta'\mathbf{Y}_t$ may have non-zero mean.

IV. Model $H^*(r)$: $\mu_t = \mu_0 + \alpha\rho_1 t$ (restricted trend). The restricted VECM is

$$\Delta \mathbf{Y}_t = \mu_0 + \alpha\left(\beta'\mathbf{Y}_{t-1} + \rho_1 t\right) + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \varepsilon_t, \tag{4.24}$$

the series in $\mathbf{Y}_t$ are $I(1)$ with drift vector μ_0 and the cointegrating relations $\beta'\mathbf{Y}_t$ have a linear trend term $\rho_1 t$.

V. Model $H(r)$: $\mu_t = \mu_0 + \mu_1 t$ (unrestricted constant and trend). The unrestricted VECM is

$$\Delta \mathbf{Y}_t = \mu_0 + \mu_1 t + \alpha\beta'\mathbf{Y}_{t-1} + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \varepsilon_t, \tag{4.25}$$

the series in $\mathbf{Y}_t$ are $I(1)$ with a linear trend (quadratic trend in levels) and the cointegrating relations $\beta'\mathbf{Y}_t$ have a linear trend.

Case I is not really relevant for empirical work. The restricted constant Case II is appropriate for non-trending $I(1)$ data like interest rates and exchange rates. The unrestricted constant Case II is appropriate for trending $I(1)$ data like asset prices, macroeconomic aggregates (real DGP, consumption, employment etc). The restricted trend case IV is also appropriate for trending $I(1)$ as in Case III. However, notice the deterministic trend in the cointegrating residual in Case IV as opposed to the stationary residual in Case III. Finally, the unrestricted trend Case V is appropriate for $I(1)$ data with a quadratic trend. An example might be nominal price data during times of extreme inflation.

4.3.4 Likelihood Ratio Tests for the Number of Cointegrating Vectors

The unrestricted cointegrating VECM (4.25) is denoted $H(r)$. The $I(1)$ model $H(r)$ can be formulated as the condition that the rank of $\mathbf{\Pi}$ is less than or equal to r. This creates a nested set of models

$$H(0) \subset \cdots \subset H(r) \subset \cdots \subset H(n)$$

where $H(0)$ represents the non-cointegrated VAR model with $\mathbf{\Pi} = \mathbf{0}$ and $H(n)$ represents an unrestricted stationary VAR(p) model. This nested formulation is convenient for developing a sequential procedure to test for the number r of cointegrating relationships.

Since the rank of the long-run impact matrix $\mathbf{\Pi}$ gives the number of cointegrating relationships in $\mathbf{Y}_t$, Johansen formulates likelihood ratio (LR) statistics for the number of cointegrating relationships as LR statistics for determining the rank of $\mathbf{\Pi}$. These test are based on the estimated eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_n$ of the matrix $\mathbf{\Pi}$. These eigenvalues also happen to equal to squared canonical correlations between $\Delta\mathbf{Y}_t$ and $\mathbf{Y}_{t-1}$ corrected for lagged $\Delta\mathbf{Y}_t$ and $\mathbf{D}_t$ and so lie between 0 and 1. Recall, the rank of $\mathbf{\Pi}$ is equal to the number of non-zero eigenvalues of $\mathbf{\Pi}$.

Johansen's Trace Statistics

Johansen's LR statistic tests the nested hypotheses

$$H_0(r): r = r_0 \text{ vs. } H_1(r): r > r_0$$

The LR statistic, called the trace statistic, is given by

$$LR_{trace}(r_0) = -T \sum_{i=r_0+1}^{n} \ln\left(1 - \hat{\lambda}_i\right)$$

If $rank(\mathbf{\Pi}) = r_0$ then $\hat{\lambda}_{r_0+1}, \ldots, \hat{\lambda}_n$ should all be close to zero and $LR_{trace}(r_0)$ should be small. In contrast, if $rank(\mathbf{\Pi}) > r_0$ then some of $\hat{\lambda}_{r_0+1}, \ldots, \hat{\lambda}_n$ will be nonzero (but less than 1) and $LR_{trace}(r_0)$ should be large. The asymptotic null distribution of $LR_{trace}(r_0)$ is not chi-square but instead is a multivariate version of Dickey-Fuller unit root distribution which depends in the dimension $n - r_0$ and the specification of the deterministic terms. Critical values for this distribution are tabulated in Osterwald-Lenum (1992) for the five trend cases discussed in the previous section for $n - r_0 = 1, \ldots, 10$.

Sequential Procedure for Determining the Number of Cointegrating Vectors

Johansen proposes a sequential testing procedure that consistently determines the number of cointegrating vectors. First test $H_0(r_0 = 0)$ against $H_1(r_0 > 0)$. If this

https://doi.org/10.26219/heal.aueb.5789

null is not rejected then it is concluded that there are no cointegrating vectors among the n variables in Y_t. If $H_0(r_0=0)$ is rejected then it is concluded that there is at least one cointegrating vector and proceed to test $H_0(r_0=1)$ against $H_1(r_0>1)$. If this null is not rejected then it is concluded that there is only one cointegrating vector. If the null is rejected then it is concluded that there is at least two cointegrating vectors. The sequential procedure is continued until the null is not rejected.

Johansen's Maximum Eigenvalue Statistic

Johansen also derives a LR statistic for the hypotheses

$$H_0(r): r = r_0 \text{ vs. } H_1(r_0): r_0 = r_0 + 1$$

The LR statistic, called the maximum eigenvalue statistic, is given by

$$LR_{\max}(r_0) = -T\ln\left(1-\hat{\lambda}_{r_0+1}\right)$$

As with the trace statistic, the asymptotic null distribution of $LR_{\max}(r_0)$ is not chi-square but instead is a complicated function of Brownian motion, which depends on the dimension $n-r_0$ and the specification of the deterministic terms. Critical values for this distribution are tabulated in Osterwald-Lenum (1992) for the five trend cases discussed in the previous section for $n-r_0=1,...,10$.

Finite Sample Correction to LR Tests

Reinsel and Ahn (1992) and Reimars (1992) have suggested that the LR tests perform better in finite samples if the factor $T-np$ is used instead of T in the construction of LR tests.

4.4 Maximum Likelihood Estimation of the Cointegrated VECM

If it is found that $rank(\mathbf{\Pi}) = r,\ 0<r<n$, then the cointegrated VECM (4.18) becomes a reduced rank multivariate regression. It is shown that

$$\hat{\beta}_{mle} = (\hat{\mathbf{v}}_1, ..., \hat{\mathbf{v}}_r), \tag{4.26}$$

where $\hat{\mathbf{v}}_i$ are the eigenvectors associated with the eigenvalues $\hat{\lambda}_i$, and that the MLEs of the remaining parameters are obtained by multivariate least squares estimation of (4.20) with β replaced by $\hat{\beta}_{mle}$.

https://doi.org/10.26219/heal.aueb.5789

Normalized Estimates

Recall, the factorization $\mathbf{\Pi} = \alpha\beta'$ is not unique and so the columns of $\hat{\beta}_{mle}$ on (4.26) may be interpreted as linear combinations of the underlying cointegrating relations. For interpretations, it is often convenient to normalize or identify the cointegrating vectors by choosing a specific coordinate system in which to express the variables. One arbitrary way to do this, suggested by Johansen, is to solve for the triangular representation of the cointegrated system. The resulting normalized cointegrating vector is denoted $\hat{\beta}_{c,mle}$. The normalization of the MLE for β to $\hat{\beta}_{c,mle}$ will affect the MLE of α but not the MLEs of the other parameters in the VECM.

It must be emphasized that it is not possible to estimate the individual elements of β without a specific normalization or identification scheme and that the normalization based on Phillips' triangular representation is arbitrary and the resulting normalized vectors $\beta_c = \begin{pmatrix} I_r \\ -\beta_1^{-1}\beta_2 \end{pmatrix}$ may mot have any economic meaning. Only in case $r = 1$ can a unique cointegrating vector be found after normalization.

https://doi.org/10.26219/heal.aueb.5789

CHAPTER 5

APPLICATION

5.1 Data

In order to visualize the usage of these techniques we are going to implement them for a five dimensional system consisting of the monthly interest rates of countries: **France, United Kingdom, Netherlands, Denmark and United States.** The data are from 1973 to 1998 in monthly frequency but we will only use data from 1974 for the analysis. The actual data are apposed at the Appendix (Table 1).

The original data have a trend and are thus considered non stationary. By using unit root test it will appear that they have a unit root so, by taking first differences, each time series will be $I(0)$ with no trend. We also extract the mean of each data series. By performing cointegration test for the set of series and by allowing linear deterministic trend, it is shown that there is long-term relationship between the series, so a VEC model will be used, and frequency domain analysis shall investigate whether there are any strong correlations between long-medium term components of data.

The spectral analysis requires all variables to be $I(0)$. A visual inspection of the data suggests that the levels of the five series are positively related and follow each other closely, see figure 5.1.

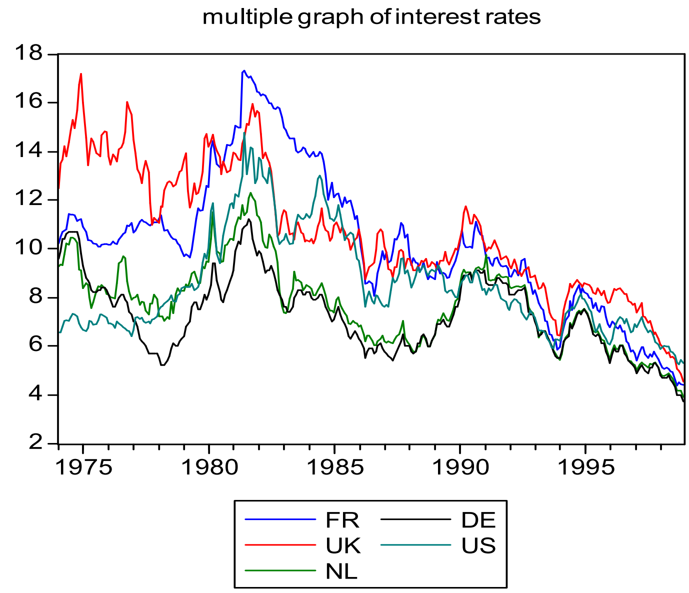


Figure 5.1. Time series of five countries interest rates

https://doi.org/10.26219/heal.aueb.5789

5.2 Unit Root Tests

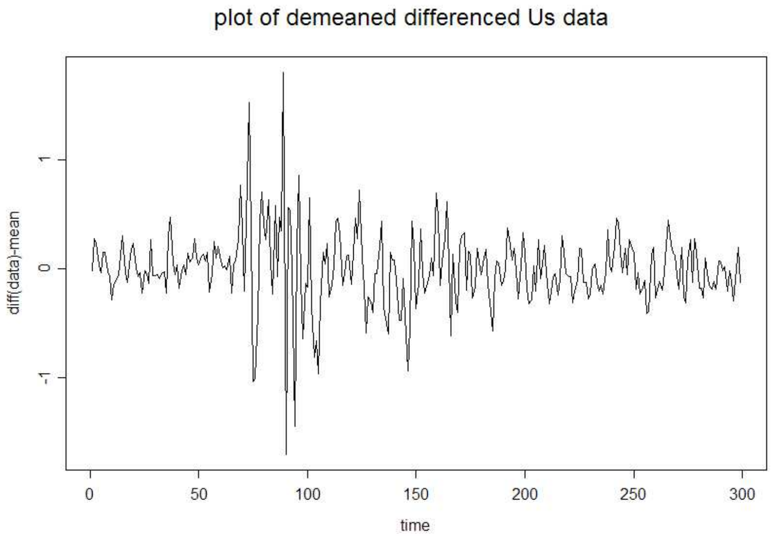
So the next thing to do is to perform unit root tests in levels for each series. The assumption made, is to include a constant and linear trend in each series. The output is apposed in Table 5.1

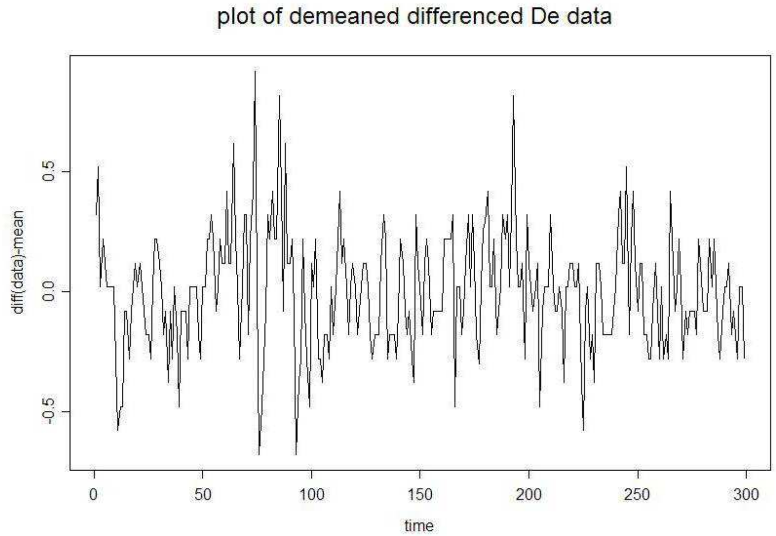
UNIT ROOT TESTS				
	LEVELS		DIFFERENCES	
COUNTRY	t-Statistic	Prob.*	t-Statistic	Prob.*
DE	2.232604	0.4692	1.096001	0.0000
FR	1.679204	0.7583	8.134046	0.0000
NL	2.357616	0.4011	1.360977	0.0000
UK	-3.812777	0.0171	-1.196192	0.0000
US	-1.682606	0.7567	-8.190398	0.0000

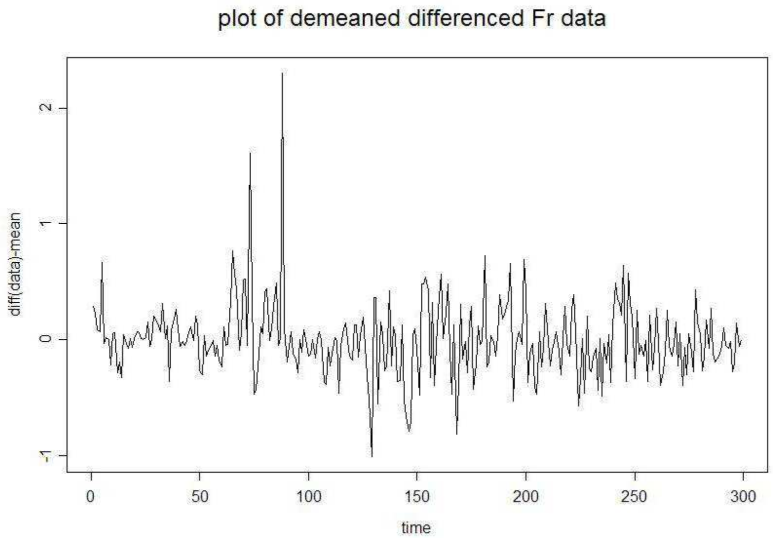
*MacKinnon (1996) one-sided p-values
Exogenous: Constant, Linear Trend
Null Hypothesis for all tests: each series has a unit root

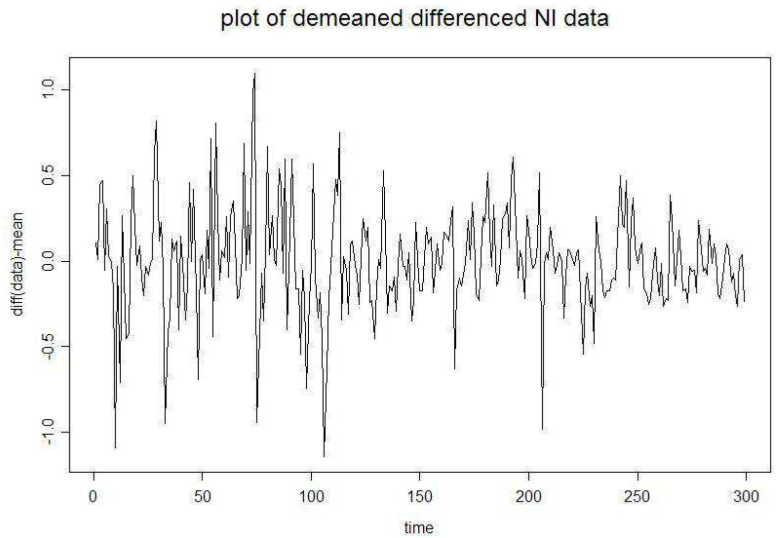
Table 5.1. Results of unit root tests

The null hypothesis of a unit root test is rejected against the one-sided alternative if the t-statistic is less than the critical value. As we can see, all tests fail to reject the null hypothesis that the series have a unit root test in levels but it exceeds rejecting the same hypothesis in first differences, fact which indicates that the data are I(1). An exception may be observed for UK's data. So when spectral analysis will be performed, the first order lagged values of the series will be used, which are represented below by dDE, dFR, dNL, dUK, and dUS. We graphically represent the differenced and demeaned data of the 5 selected countries, see figure 5.2.


plot of demeaned differenced Us data
diff(data)-mean
time


plot of demeaned differenced De data
diff(data)-mean
time


plot of demeaned differenced Fr data
diff(data)-mean
time


plot of demeaned differenced Nl data
diff(data)-mean
time

https://doi.org/10.26219/heal.aueb.5789

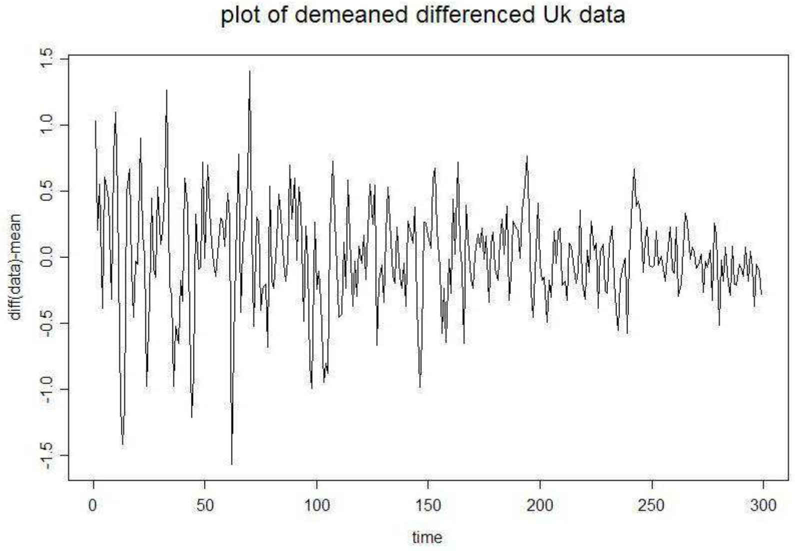


Figure 5.2. Plots of demeaned differenced data of each time series

5.3 Cointegration Analysis of Interest Rates

Moreover we have to speculate whether there exists a long term relationship between the series. When this kind of relationship exists, is coped with a VEC model, which is a restricted VAR that has cointegration restrictions built into the specification, so it is designed for use of nonstationary series that are known to be cointegrated. The VEC specification restricts the long run behaviour of the endogenous variables to converge to their cointegrating relationships while allowing a wide range of short-run dynamics.

Sample(adjusted): 1974:04 1998:12
Included observations: 297 after adjusting endpoints
Trend assumption: Linear deterministic trend
Series: DE FR NL UK US
Lags interval (in first differences): 1 to 2
Unrestricted Cointegration Rank Test

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	5 Percent Critical Value	1 Percent Critical Value
None *	0.092977	72.35412	68.52	76.07
At most 1	0.073386	43.37051	47.21	54.46
At most 2	0.044446	20.73356	29.68	35.65
At most 3	0.022923	7.230897	15.41	20.04
At most 4	0.001156	0.343393	3.76	6.65

*(**) denotes rejection of the hypothesis at the 5%(1%) level
Trace test indicates 1 cointegrating equation(s) at the 5% level
Trace test indicates no cointegration at the 1% level

Table 5.2. Cointegration tests for the set of series

https://doi.org/10.26219/heal.aueb.5789

The cointegration test, where results are summarized in Table 5.2, indicates that 1 cointegrating equation exists at 5% but does not hold at 1% level, so long term relationship between the countries is proved, as Juselius had maintained at her work. An attempt to clarify which of the five countries take part mostly in this relationship, by restricting the VEC model, shows that NL, DE, and UK are those countries who participate forcefully in the equation as shown below. This fact is determined again by the minimized AIC criterion, in Table 5.3.

Vector Error Correction Estimates
Date: 05/28/07 Time: 13:24
Sample(adjusted): 1974:04 1998:12
Included observations: 297 after adjusting endpoints
Standard errors in ( ) & t-statistics in [ ]

Cointegration Restrictions:
B(1,1)=0,B(1,5)=0
Convergence achieved after 6 iterations.
Not all cointegrating vectors are identified
LR test for binding restrictions (rank = 1):

Chi-square(2)	2.310410
Probability	0.314993

Cointegrating Eq:	CointEq1
FR(-1)	0.000000
UK(-1)	-0.473582
NL(-1)	1.665657
DE(-1)	-1.165727
US(-1)	0.000000
C	0.645605

Determinant Residual Covariance	1.79E-06
Log Likelihood	-111.2310
Log Likelihood (d.f. adjusted)	-141.8539
Akaike Information Criteria	1.392956
Schwarz Criteria	2.201348

Table 5.3. Results of restricted VEC model

For forecasting this system of interrelated interest rates series and for analysing the dynamic impact of random disturbances on the systems variables we will use the VEC model method. The VEC approach sidesteps the need for structural modelling by modelling every endogenous variable in the system (which is the time series of every country's interest rate) as a function of the lagged values of all of the endogenous variables in the system and also takes concern of the cointegrated relationships of the series. The number of the lags included in the model is determined by the smaller value of the AIC criteria. The smallest value seems to correspond to VEC(2) model. So the estimated values of each interest rate are modelled by:

dFR = $0.009187134074*(- 0.4735818587*UK_{t-1} + 1.665656873*NL_{t-1} - 1.165726628*DE_{t-1} + 0.6456053157) + 0.1292183664* dFR_{t-1} - 0.1017579722* dFR_{t-2} + 0.005113179693* dUK_{t-1} - 0.0183230216* dUK_{t-2} + 0.01166026533* dNL_{t-1} + 0.0737411488* dNL_{t-2} - 0.08640053961* dDE_{t-1} + 0.1512769108* dDE_{t-2} + 0.1899385723* dUS_{t-1} - 0.001383204304* dUS_{t-2} - 0.0174686818$

dUK = $0.03346411087*(- 0.4735818587*UK_{t-1} + 1.665656873*NL_{t-1} - 1.165726628*DE_{t-1} + 0.6456053157) + 0.02074353747* dFR_{t-1} + 0.05826566062* dFR_{t-2} + 0.3065960726* dUK_{t-1} - 0.1574120191* dUK_{t-2} - 0.1006058681* dNL_{t-1} + 0.137973906* dNL_{t-2} + 0.2205150912* dDE_{t-1} + 0.01078362491* dDE_{t-2} + 0.119010806* dUS_{t-1} - 0.1324505776* dUS_{t-2} - 0.01923052814$

dNL = $- 0.04993762697*(- 0.4735818587*UK_{t-1} + 1.665656873*NL_{t-1} - 1.165726628*DE_{t-1} + 0.6456053157) + 0.05665107092* dFR_{t-1} - 0.08764156779* dFR_{t-2} - 0.04011022607* dUK_{t-1} - 0.01497490525* dUK_{t-2} + 0.06788147222* dNL_{t-1} - 0.03435194143* dNL_{t-2} + 0.2541409734* dDE_{t-1} + 0.04979691941* dDE_{t-2} + 0.1848312434* dUS_{t-1} - 0.01375094647* dUS_{t-2} - 0.01314811169$

dDE = $0.02497211711*(- 0.4735818587*UK_{t-1} + 1.665656873*NL_{t-1} - 1.165726628*DE_{t-1} + 0.6456053157) + 0.04351164725* dFR_{t-1} - 0.04908308961* dFR_{t-2} + 0.03608812295* dUK_{t-1} - 0.03578546719* dUK_{t-2} + 0.168248814* dNL_{t-1} - 0.07019409355* dNL_{t-2} + 0.1736228788* dDE_{t-1} + 0.03063968151* dDE_{t-2} + 0.153123193* dUS_{t-1} + 0.007567419831* dUS_{t-2} - 0.01611017716$

dUS = $0.004038567652*(- 0.4735818587*UK_{t-1} + 1.665656873*NL_{t-1} - 1.165726628*DE_{t-1} + 0.6456053157) + 0.2268441094* dFR_{t-1} - 0.131181081* dFR_{t-2} + 0.01790620552* dUK_{t-1} + 0.01208104519* dUK_{t-2} + 0.1327248652* dNL_{t-1} + 0.001984167895* dNL_{t-2} + 0.05057755* dDE_{t-1} - 0.1967101223* dDE_{t-2} + 0.1367690701* dUS_{t-1} - 0.1388051532* dUS_{t-2} - 0.002648186992$

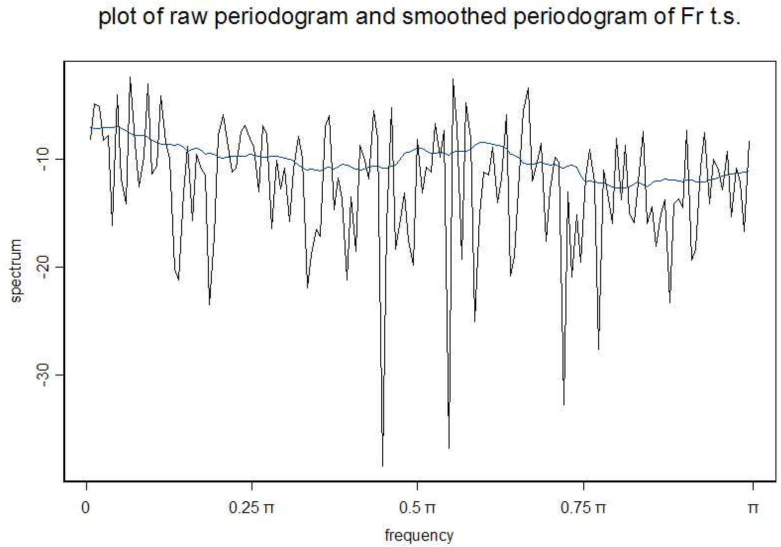
https://doi.org/10.26219/heal.aueb.5789

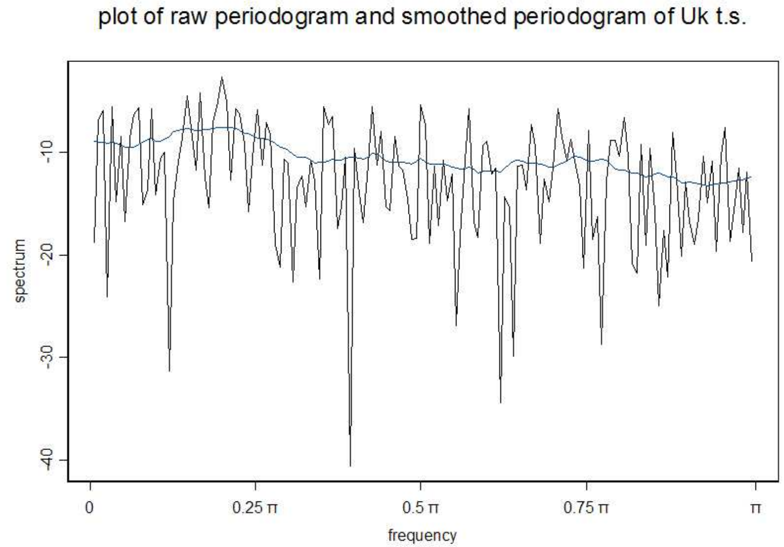
5.4 Multivariare frequency domain analysis

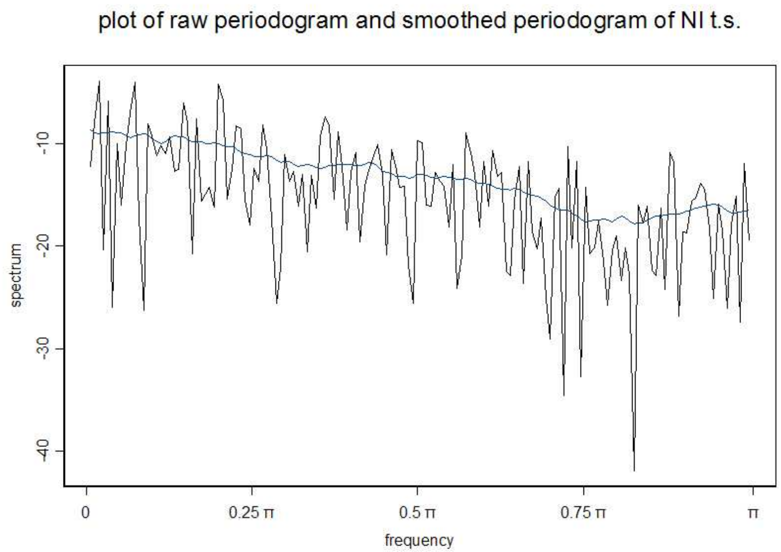
5.4.1 Individual spectra

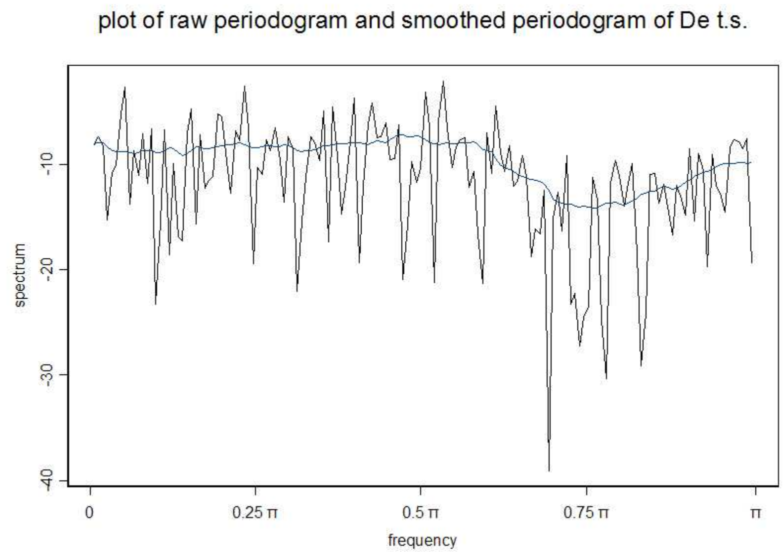
The non-parametric estimates of the spectrum, coherency and phase are available for estimating the spectrum and for revealing features of the data like the correlation between series due to frequency size (low, medium, high). The autoregressive procedure however, as Ioannidis *(2007)* notes, has a much more clearly defined criterion for selecting *m* (the order of the model) than the corresponding criteria to be considered in the selection of a spectral window. In estimating a spectral density it is wise to examine both types of density estimators. Parzen *(1974)* has also suggested that the cumulative periodogram should be compared with the autoregressive estimate of the spectral distribution function as an aid to autoregressive model selection in the time domain.

So the next thing to do is to present graphically the raw periodogram of the series combined with their smoothed periodogram.


plot of raw periodogram and smoothed periodogram of Fr t.s.
spectrum
-10
-20
-30
0
0.25 π
0.5 π
0.75 π
π
frequency


plot of raw periodogram and smoothed periodogram of Uk t.s.
spectrum
-10
-20
-30
-40
0
0.25 π
0.5 π
0.75 π
π
frequency


plot of raw periodogram and smoothed periodogram of Nl t.s.
spectrum
-10
-20
-30
-40
0
0.25 π
0.5 π
0.75 π
π
frequency


plot of raw periodogram and smoothed periodogram of De t.s.
spectrum
-10
-20
-30
-40
0
0.25 π
0.5 π
0.75 π
π
frequency

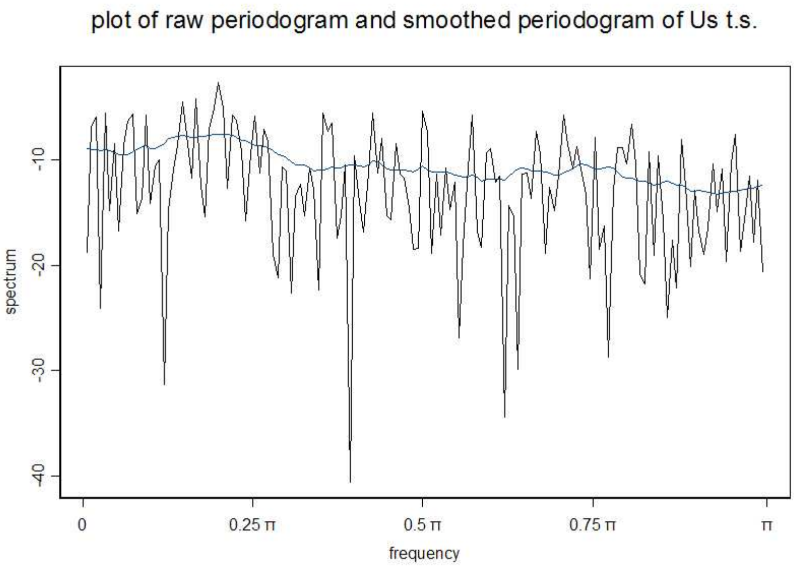
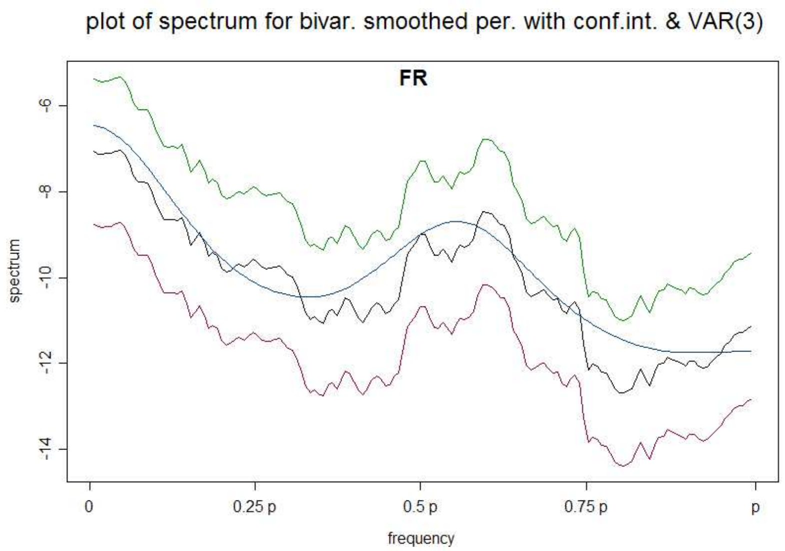


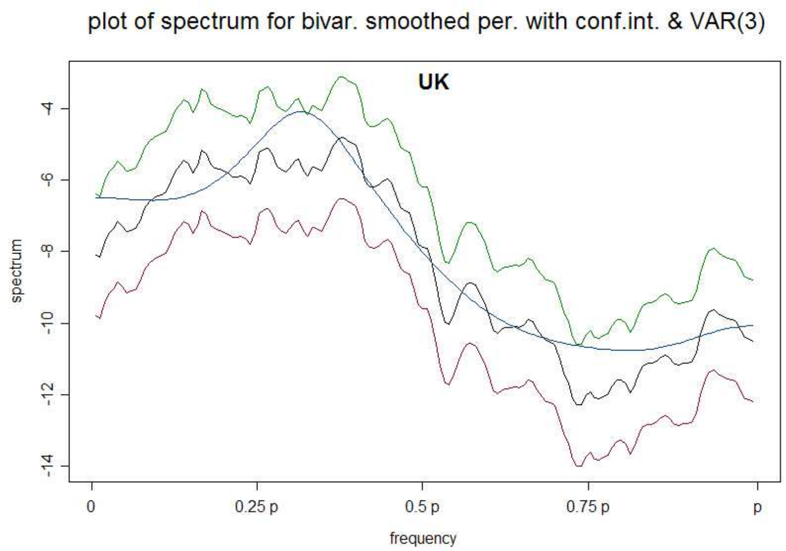
Figure 5.3. Plots of raw and smoothed periodogram of each time series

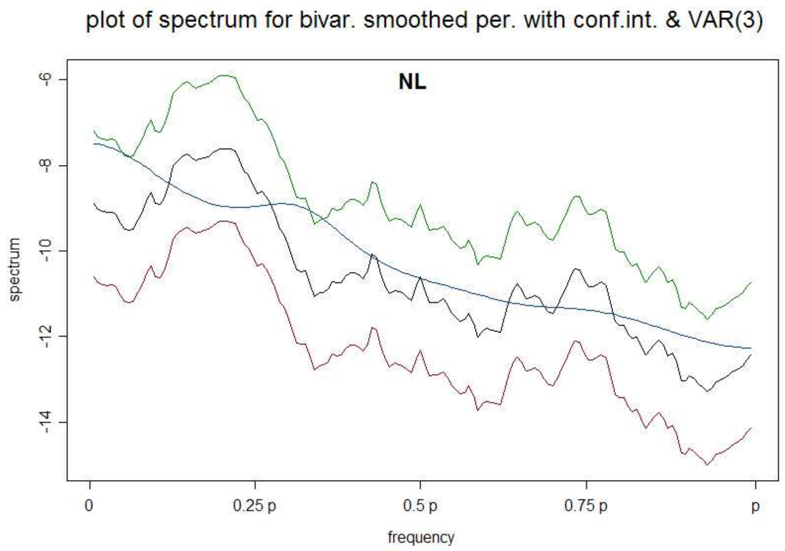
The smoothness of the periodogram of all series is obtained by using the Daniell window, which corresponds to the discrete spectral average estimator with weights $W_n(j) = (2m+1)^{-1}$, $|j| \leq m = [r/2n]$ with m=12. As we can see, all sharp peacks have been smoothed. It is noticeable that all periodograms are rather smooth, with no mentionable extreme values to present. This picture, in association with the fact that lines represent the differenced series, clues that the original data are I(1). More specific, they appear high values at low frequencies, gradually decreased at high frequencies, (close to zero) especially near $\lambda = 0.75\pi$, a number which corresponds to 3-month waves.

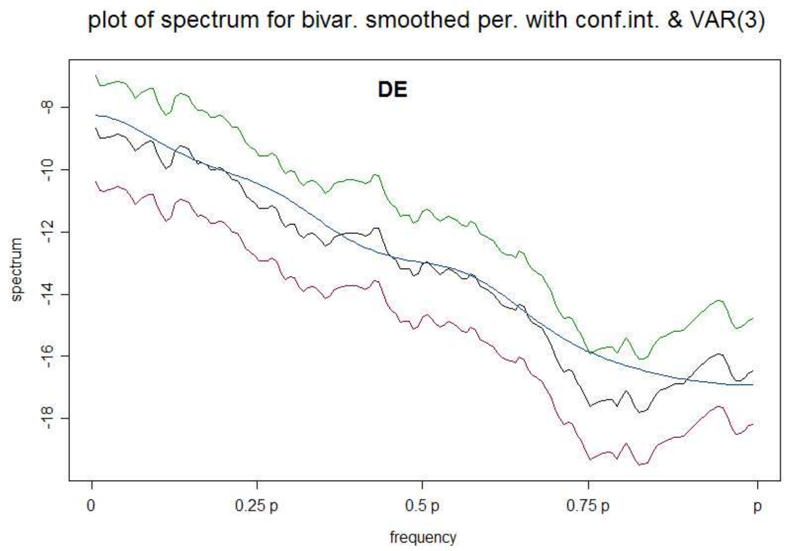
Next we try to fit a multivariate autoregressive model according to the minimized AIC criterion. As shown in Table 2 of the Appendix. the VAR(3) model is chosen. The graphical presentation of the smoothed periodogram with its confidence intervals versus the parametric Var(3) spectrum is presented below. The smoothed periodogram reveals no significant periodicities except from a minimum value at π/2. As we can see the parametric model suits quite well to the non parametric spectrum of all series and catches extreme values successfully in almost all cases except Netherlands. As we can see for spectra of UK, DE, NL, and US the power is mostly concentrated in low frequencies.

.


plot of spectrum for bivar. smoothed per. with conf.int. & VAR(3)
FR
spectrum
frequency
0
0.25 p
0.5 p
0.75 p
p


plot of spectrum for bivar. smoothed per. with conf.int. & VAR(3)
UK
spectrum
frequency
0
0.25 p
0.5 p
0.75 p
p


plot of spectrum for bivar. smoothed per. with conf.int. & VAR(3)
NL
spectrum
frequency
0
0.25 p
0.5 p
0.75 p
p


plot of spectrum for bivar. smoothed per. with conf.int. & VAR(3)
DE
spectrum
frequency
0
0.25 p
0.5 p
0.75 p
p

 https://doi.org/10.26219/heal.aueb.5789

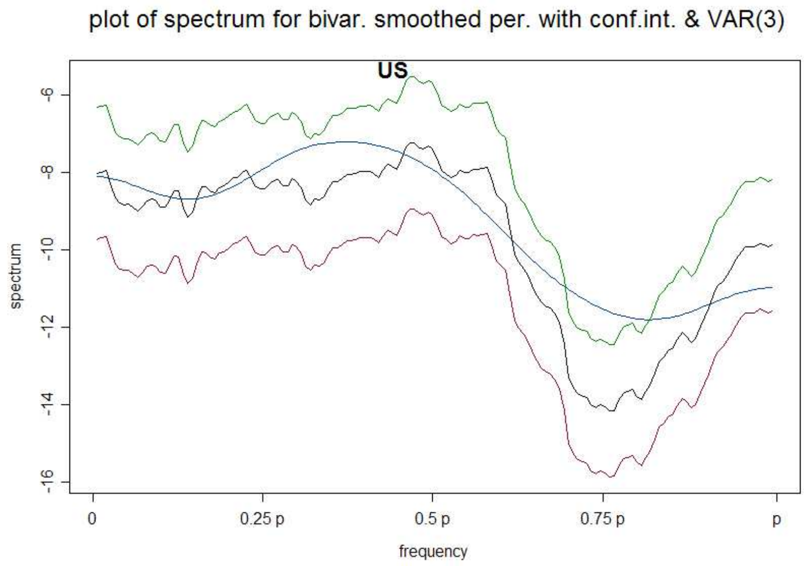


Figure 5.4. Spectrum with confidence intervals due to VAR(3) parametric model of each time series

5.4.2 Coherency-Phase between Interest Rates

At this section we will work with the coherencies and the phases between each pair of time series (which means 10 pairs).

Regarding the coherencies we notice a common behavior for almost all of the series at low frequencies. That is, really high values (from 0.6 till 0.9) for frequencies from zero up to π/4 (8 month relationship). Specially the pairs: NL-DE, UK-NL, UK-DE seem to be nearly linearly correlated at (0.8, 0.9) frequency interval. On the other hand the FR-UK pair is less coherent than the others, with a coherency of about 0.55. This remark is the same with Juselius conclusion that British, Spanish, Danish and German rates have a common behavior and Italian, Dutch, Belgian and French bond rate belong to a group of different behavior. Coherencies for medium and short term fluctuations are generally lower than for long term fluctuations, although still not negligible.

At medium frequencies, from π/4 till 0.75π, for the pairs NL-US, DE-US and NL-DE, coherency falls gradually and at 0.75π frequency exceeds its lowest value, near zero. This means that for 3 month period fluctuations there is no influence between these countries. The rest of the countries begin with a low value near after π/4, which grows till π/2 where it exceeds a high value and falls off till 0.75π. For the rest of the frequency interval, behaviors vary a lot between the subsets of countries. Coherency of FR-UK, FR-US, increases at this interval and near π exceeds nearly the same value as that of π/2 frequency. On the other hand coherency of FR-DE, NL-DE, falls gradually to zero. For the rest of them, FR-NL, UK-DE, UK-NL, DE-US, UK-US, coherency falls near zero shortly after 0.75π, increases till 0.9π and decreases for the rest of the period. As a total behavior we have to underline that the pairs NL-US, DE-US and mostly NL-DE have really high correlation at low frequencies and decrease till zero at π. More generally they are close to 0 for low frequencies, meaning that long term fluctuations are synchronous across countries.

In some other pairs, coherency drops to zero irregularly in all range of frequencies. From economic viewpoint this particularly means that some countries act almost independent in high frequencies and on the other hand seem to have strong coherent behaviour, and almost linear in low frequencies, for example the NL-US or the NL-De interest rates.

Phase graphs have a more common behaviour between pairs of countries. One kind of behaviour, pictures phases near zero for all frequencies until 3π/4, and concludes with high values at π. This pattern is represented by UK-NL, NL-US, FR-US. The opposite behavior, phase near zero until 3π/4 and negative values for the rest of the period present UK-US, DE-US, UK-DE. The rest of pairs fluctuate around zero but the phase's confidence intervals are wide open for the corresponding frequencies where their coherency lies near zero. Summarizing the above, coherency at low frequencies exceeds high values, in some cases is nearly one, almost linearly related, but decreases at high frequencies. Each pair's spectral coherency versus phase diagram indicates that, when coherency converges to zero, then naturally the phase distributes uniformly. The frequency interval, where the series have this behaviour, is

$$\left[\frac{3\pi}{4},\pi\right]$$

The coherency and the phase lines of the VAR fit model is also drawn. Although the parametric model seems to fit quite well for the spectrum, the inspection of the non parametric coherency due to the parametric coherency frequently slips from the non parametric confidence intervals for some of cases. For the pairs FR-UK, FR-NL, FR-US, NL-US, DE-US, UK-US, UK-NL the parametric coherency lies between the non parametric confidence intervals. But if we examine the coherences of UK-DE, NL-DE and FR-DE, we can see that parametric coherency has totally different behavior from the non parametric one. For the last set of countries the same

https://doi.org/10.26219/heal.aueb.5789

disparity appears for the parametric and non parametric phase. Although the order of the VAR model may work for fitting well the estimated spectrum of each series, the same order is possible to be inadequate for estimating the measure of coherence of each pair of series or for emulating its phase. In our group of series the VAR(3) model estimates efficiently each spectrum but in the same time, especially for pairs of DE with the other countries, diverges from the confidence intervals of the spectral coherency. This fact does not necessary mean that the order of the VAR model is inadequate and that if we assume a higher order VAR model then the coherency's divergence would be corrected.

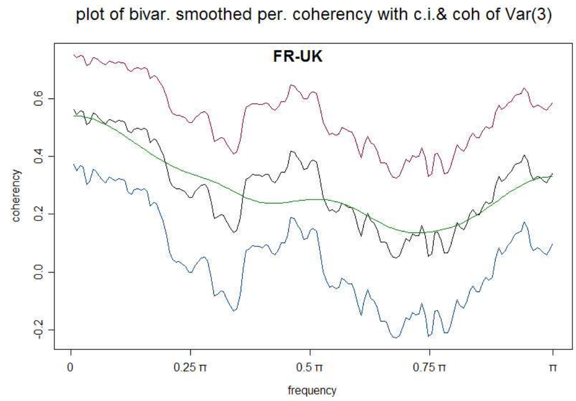
It is possible that a VAR(p) model is inadequate for the differenced data. There are two indications in this direction:

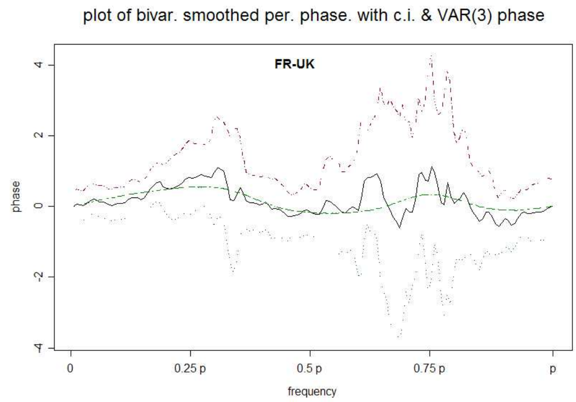
1. Even if the levels follow a VAR(p) model, if there is cointegration, theory tells us that the differenced data will not follow a VAR(∞). In our case, we confirm one cointegration relationship however, coherencies close to 1 indicate that we might be close to such a situation.

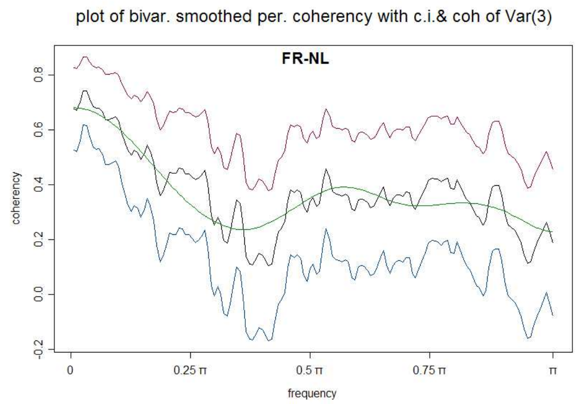
2. In a VAR(p) model coherencies close to 1 at $\lambda=0$ imply forcefully peaks of the individual spectra at $\lambda=0$. In our case coherencies close to 1 coexist with marginal spectra which are close to white noise. This is not possible in a VAR(p) model.

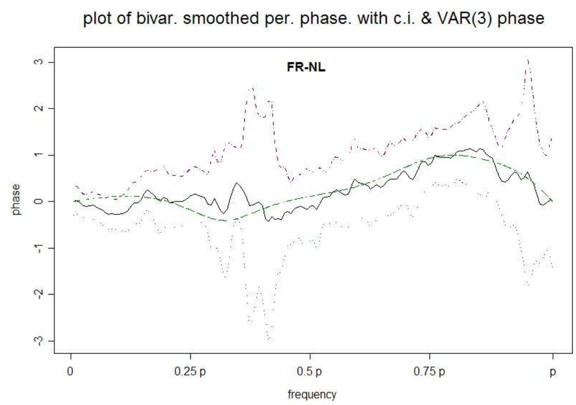
The coherency and the phase of each pair of countries is apposed below.

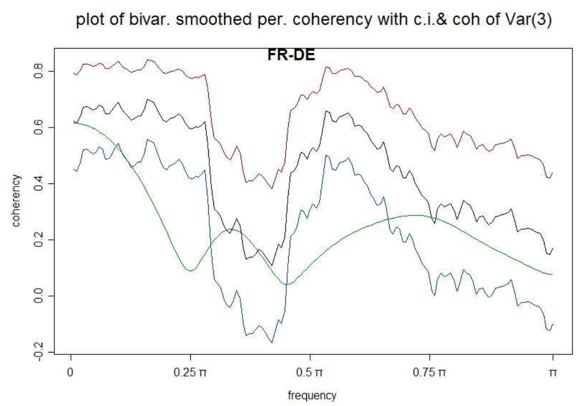
https://doi.org/10.26219/heal.aueb.5789

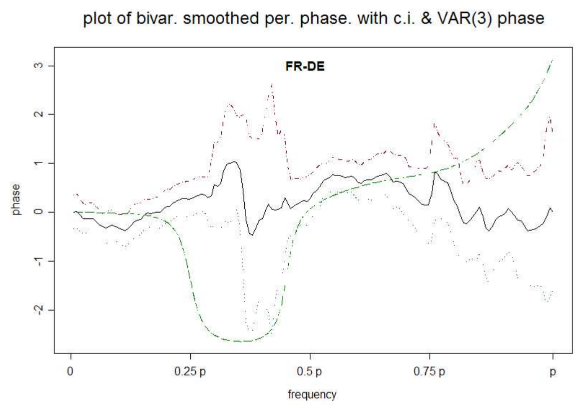


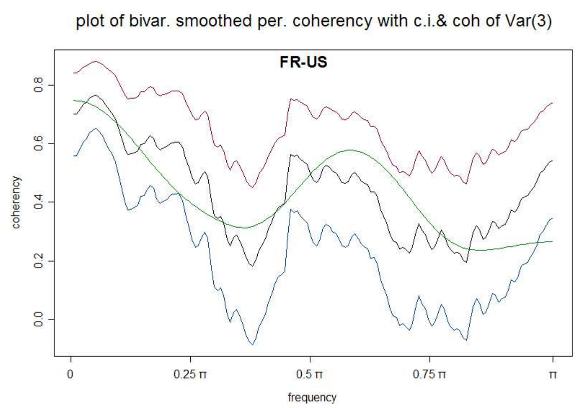


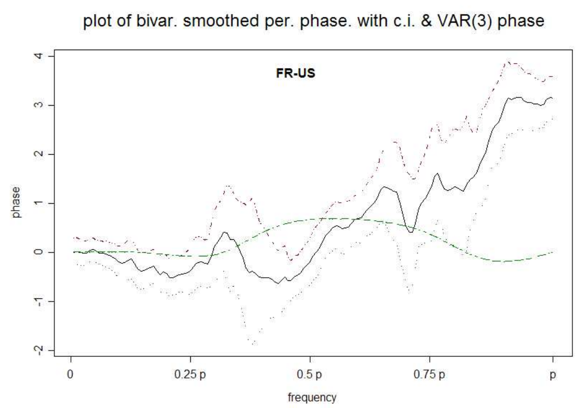





plot of bivar. smoothed per. coherency with c.i.& coh of Var(3)
FR-DE
coherency
-0.2
0.0
0.2
0.4
0.6
0.8
0
0.25 π
0.5 π
0.75 π
π
frequency

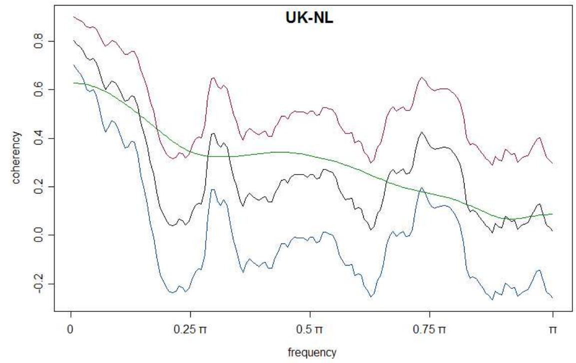

plot of bivar. smoothed per. phase. with c.i. & VAR(3) phase
FR-DE
phase
-2
-1
0
1
2
3
0
0.25 p
0.5 p
0.75 p
p
frequency


plot of bivar. smoothed per. coherency with c.i.& coh of Var(3)
FR-US
coherency
0.0
0.2
0.4
0.6
0.8
0
0.25 π
0.5 π
0.75 π
π
frequency

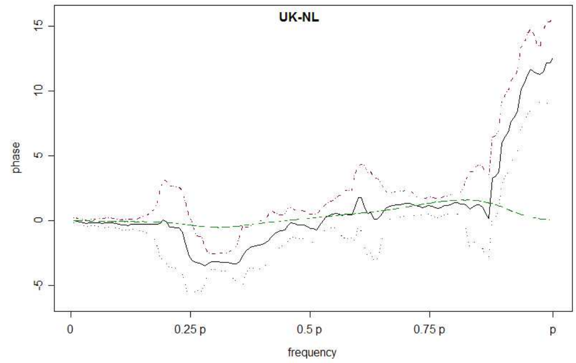

plot of bivar. smoothed per. phase. with c.i. & VAR(3) phase
FR-US
phase
-2
-1
0
1
2
3
4
0
0.25 p
0.5 p
0.75 p
p
frequency

https://doi.org/10.26219/heal.aueb.5789

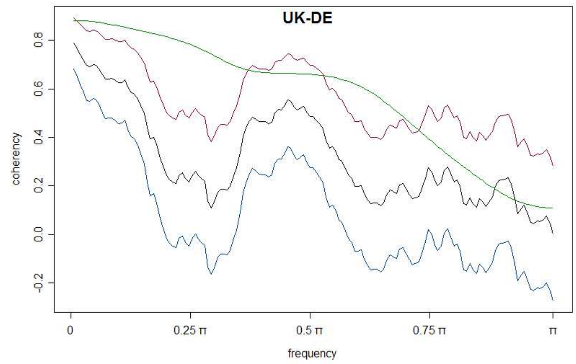
plot of bivar. smoothed per. coherency with c.i.& coh of Var(3)



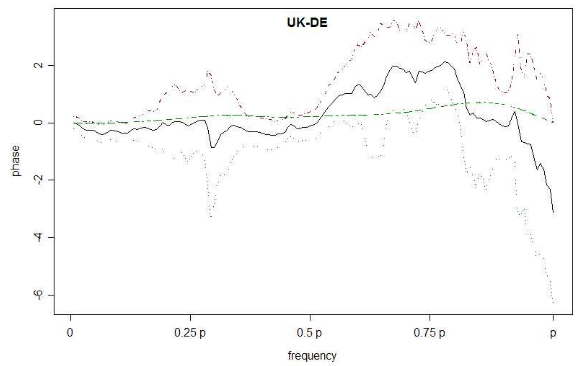
plot of bivar. smoothed per. phase. with c.i. & VAR(3) phase



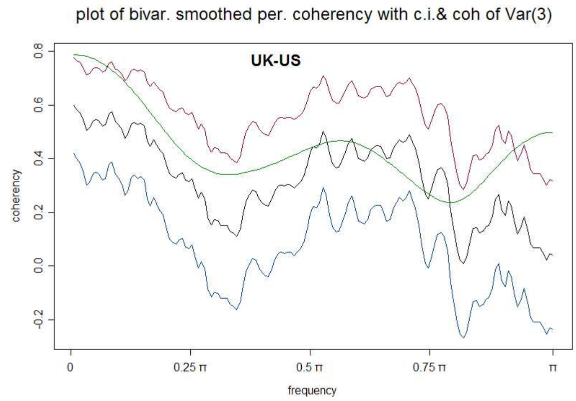
plot of bivar. smoothed per. coherency with c.i.& coh of Var(3)

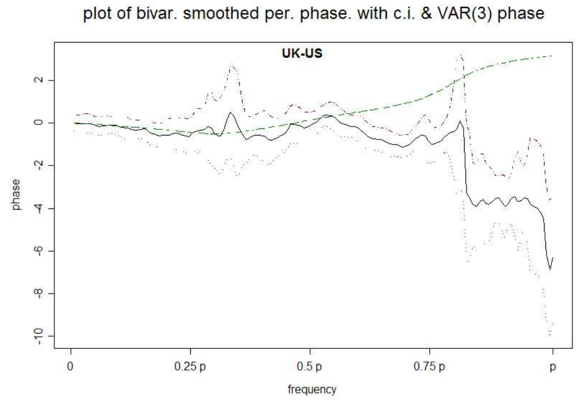


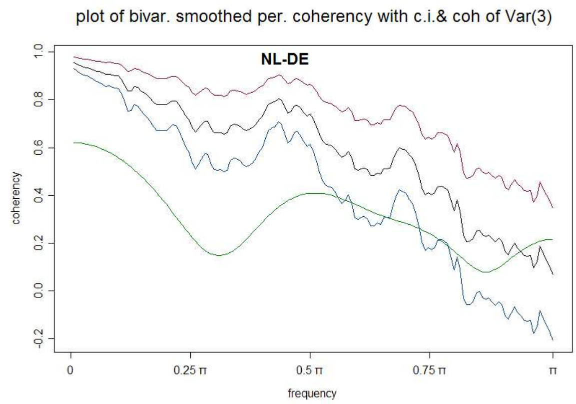
plot of bivar. smoothed per. phase. with c.i. & VAR(3) phase

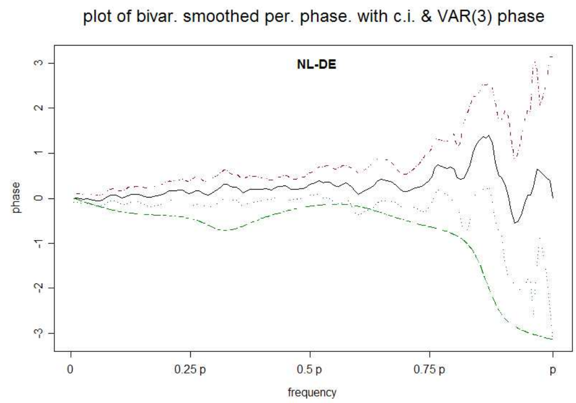


https://doi.org/10.26219/heal.aueb.5789



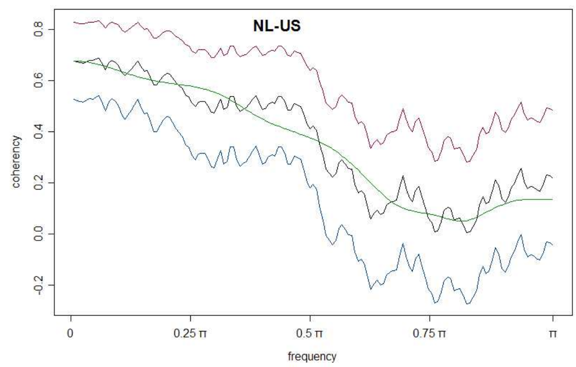




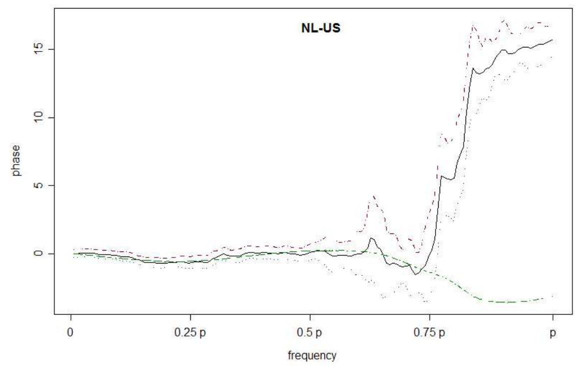


https://doi.org/10.26219/heal.aueb.5789

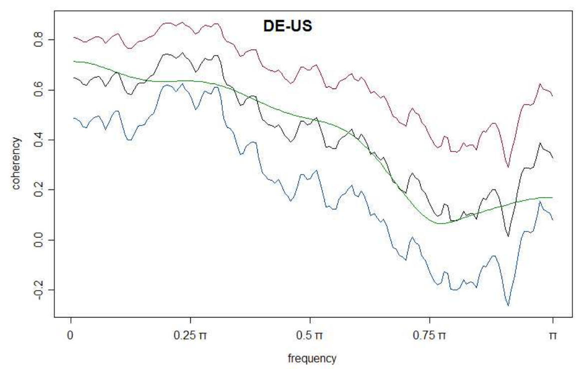
plot of bivar. smoothed per. coherency with c.i.& coh of Var(3)



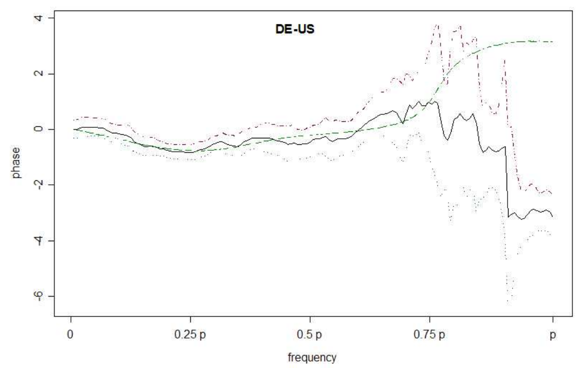
plot of bivar. smoothed per. phase. with c.i. & VAR(3) phase



plot of bivar. smoothed per. coherency with c.i.& coh of Var(3)



plot of bivar. smoothed per. phase. with c.i. & VAR(3) phase



https://doi.org/10.26219/heal.aueb.5789

Figure 5.5. i) Coherency with its confidence intervals due to parametric model coherency for each combination of countries and
ii) phase with its confidence intervals due to parametric model phase for each combination of countries

CHAPTER 6

CONCLUSIONS

After studies based on a VAR model and inquiring for cointegration relationships between the countries, Juselius derived interesting conclusions. Firstly she determined cointegration relationship between six of the nine series of the set of countries. The period investigated was relatively calm so maybe the conclusions do not typify extraordinary events and nervous markets. The presence of this relationship causes related behavior between the countries. The German, Belgian and Dutch bond rates exclusively react on the lagged change in US bond rate, but not on lagged changes of any of the European bond rates. The Spanish, Danish, French and Italian bond rates seem to have been more responsive to lagged shocks to the European bond rates. The lagged change of US bond rate is significantly affecting all the European bond rates whereas none of the European bond rates are important for US bond rate. Concluding she noticed that Italian, Dutch, Belgian and French bond rate seemed to be primarily pushing, whereas British, Spanish, Danish and German were found to be primarily adjusting. This seemed to point to the relative importance of the demand and supply of government bonds and, hence government debt, as an explanation to stochastic trends in the Euro bond rates.

Our main conclusions, derived from the analysis of monthly data for the same set of countries, are:

- There exists one cointegration relationship between the series, so there is a long-term relationship between the series. The information of which countries take part mostly in this relationship is derived by restricting the VEC model and shows that NL, DE, and UK are those countries who participate forcefully in the equation.

- For frequencies from zero up to π/4 (8 month relationship), pairs NL-DE, UK-NL, UK-DE have coherencies with very high values (from 0.8 to 1), i.e. long-medium IR fluctuations for these pairs of countries are very strongly correlated

- At medium frequencies, coherency of most pairs of countries fluctuates, near π/2, it takes values near zero, which shows the independence of the series at 4 month interval, i.e. medium term IR fluctuations tend to be uncorrelated cross countries.

- At high frequencies (0.75π, π), the coherence's behavior varies between pairs. Some of them increase and some of them gradually fall till near zero.

- The phase for most of pairs fluctuates near zero for low and medium frequencies. This means that long and medium term IR fluctuations are synchronous across countries. For high frequencies, some pairs have phase which exceeds high values, some others drop to negative values. But for such frequencies confidence intervals are very wide, indicating a high degree of uncertainty for this behavior.

- The multivariate autoregressive model selected by the minimized AIC criterion is of order 3. This is means that, in order to estimate future values for a series, 3 times above observations of each series, are effectual for a valid estimation.

As we can see, the conclusions derived from the daily data are similar to those derived from the monthly data.

Although Juselius (2001) in her daily analysis of daily data finds three cointegrating relations among European countries and the United States in this analysis of monthly data we only find one cointegrating relationship. This is not surprising since the time horizons are different: adjustment mechanisms and dynamics which may be visible and detectable on a day-to-day basis may vanish in a month-to-month basis. Moreover processes that may appear stationary on a month-to-month basis may appear non-stationary on a day-to-day basis. Finally, some discrepancies between our results and the ones by Juselius are surely due to the fact that in our analysis Spain, Denmark, Belgium and Italy are not included in the analysis. There is, however, one striking fact in which the two analysis coincide: the cointegration relation found in the present thesis concerns the same countries which are present in the first cointegration relation of Juselius (beyond Spain and Denmark which are not included in our analysis): Germany, the United Kingdom and the Netherlands.

.

https://doi.org/10.26219/heal.aueb.5789

APPENDIX

Table 1 (Data Set from Jan. 1974 to Dec. 1998)

Time	fr	Uk	nl	de	us
1974m01	10.21	12.50	9.26	9.60	6.56
1974m02	10.48	13.50	9.35	9.90	6.54
1974m03	10.69	13.68	9.34	10.40	6.81
1974m04	10.75	14.21	9.77	10.40	7.04
1974m05	10.80	13.80	10.22	10.60	7.07
1974m06	11.45	14.38	10.15	10.70	7.03
1974m07	11.40	14.88	10.44	10.70	7.18
1974m08	11.40	15.29	10.45	10.70	7.33
1974m09	11.38	14.95	10.43	10.70	7.30
1974m10	11.14	15.68	10.25	10.70	7.22
1974m11	11.17	16.75	9.14	10.40	6.93
1974m12	11.21	17.18	9.09	9.80	6.78
1975m01	10.90	16.02	8.36	9.30	6.68
1975m02	10.69	14.58	8.61	8.80	6.61
1975m03	10.34	13.43	8.49	8.70	6.73
1975m04	10.36	13.89	8.02	8.60	7.03
1975m05	10.31	14.53	7.58	8.30	6.99
1975m06	10.21	14.41	7.80	8.20	6.86
1975m07	10.20	13.93	8.28	8.20	6.89
1975m08	10.11	13.87	8.47	8.30	7.06
1975m09	10.10	13.79	8.43	8.30	7.29
1975m10	10.15	14.66	8.50	8.40	7.29
1975m11	10.19	14.81	8.38	8.40	7.21
1975m12	10.18	14.79	8.16	8.30	7.17
1976m01	10.16	13.79	8.11	8.10	6.94
1976m02	10.16	13.46	8.01	7.90	6.92
1976m03	10.29	13.88	7.98	7.60	6.87
1976m04	10.21	13.77	7.97	7.60	6.73
1976m05	10.19	13.59	8.58	7.80	6.99
1976m06	10.37	14.09	9.38	8.00	6.92
1976m07	10.51	14.16	9.48	8.10	6.85
1976m08	10.61	14.33	9.69	8.10	6.79
1976m09	10.66	14.79	9.63	7.90	6.70
1976m10	10.95	16.03	8.66	7.80	6.65
1976m11	10.94	15.79	8.23	7.40	6.62
1976m12	11.04	15.48	7.88	7.30	6.39

https://doi.org/10.26219/heal.aueb.5789

1977m01	10.66	14.48	7.99	7.00	6.68
1977m02	10.72	13.93	8.03	7.00	7.15
1977m03	10.88	13.25	8.13	6.80	7.20
1977m04	11.12	13.05	7.71	6.30	7.14
1977m05	11.19	12.69	7.84	6.20	7.17
1977m06	11.11	13.26	7.81	6.10	6.99
1977m07	11.07	13.62	7.45	6.00	6.97
1977m08	11.00	13.12	7.30	5.70	7.00
1977m09	10.97	11.88	7.74	5.70	6.94
1977m10	11.01	10.98	7.72	5.70	7.08
1977m11	11.10	11.28	8.12	5.70	7.14
1977m12	11.07	11.16	7.93	5.70	7.23
1978m01	11.25	11.06	7.22	5.50	7.50
1978m02	11.36	11.75	7.21	5.20	7.60
1978m03	11.07	11.72	7.23	5.20	7.63
1978m04	10.75	12.39	7.02	5.20	7.74
1978m05	10.76	12.72	7.18	5.40	7.87
1978m06	10.60	12.79	7.12	5.60	7.94
1978m07	10.49	12.72	7.82	5.90	8.09
1978m08	10.40	12.55	7.36	6.10	7.87
1978m09	10.37	12.64	8.15	6.00	7.82
1978m10	10.21	12.91	8.38	6.00	8.07
1978m11	10.14	13.16	8.25	6.20	8.16
1978m12	9.94	13.22	8.29	6.30	8.36
1979m01	9.68	13.68	8.29	6.40	8.43
1979m02	9.77	13.94	8.53	6.80	8.43
1979m03	9.70	12.35	8.42	6.90	8.45
1979m04	9.64	11.68	8.66	7.00	8.44
1979m05	9.96	11.94	8.99	7.60	8.55
1979m06	10.71	12.69	9.07	7.80	8.32
1979m07	11.25	12.25	8.83	7.80	8.35
1979m08	11.67	12.30	8.61	7.50	8.42
1979m09	11.56	12.60	8.61	7.50	8.68
1979m10	11.60	13.16	9.28	7.80	9.44
1979m11	12.09	14.54	9.21	8.10	9.80
1979m12	12.59	14.72	9.48	7.90	9.59
1980m01	12.52	14.17	9.45	8.10	10.03
1980m02	14.11	14.45	10.41	8.50	11.55
1980m03	14.44	14.70	11.49	9.40	11.87
1980m04	13.95	14.27	10.53	9.40	10.83
1980m05	13.49	14.01	9.91	8.70	9.82
1980m06	13.34	13.78	9.82	8.20	9.40

https://doi.org/10.26219/heal.aueb.5789

1980m07	13.43	13.07	9.45	7.90	9.83
1980m08	13.46	13.58	9.49	7.80	10.53
1980m09	13.85	13.38	10.14	8.10	10.94
1980m10	14.27	13.12	10.16	8.30	11.20
1980m11	14.24	13.22	10.41	8.70	11.83
1980m12	14.31	13.67	10.40	8.90	11.89
1981m01	14.58	13.96	10.36	9.10	11.65
1981m02	15.05	13.89	10.88	9.90	12.23
1981m03	14.98	13.68	11.31	10.30	12.15
1981m04	14.97	13.64	11.22	10.20	12.62
1981m05	17.25	14.31	11.80	10.80	12.96
1981m06	17.32	14.57	11.38	10.90	14.76
1981m07	17.11	15.14	11.50	11.00	13.05
1981m08	17.03	15.09	12.08	11.20	13.61
1981m09	17.08	15.59	12.30	11.10	14.14
1981m10	16.94	15.95	12.12	10.40	14.13
1981m11	16.75	15.44	11.94	10.00	12.68
1981m12	16.44	15.65	11.38	9.70	12.88
1982m01	16.42	15.58	11.31	9.90	13.73
1982m02	16.29	14.74	11.07	9.80	13.63
1982m03	16.35	13.72	10.31	9.50	12.98
1982m04	16.29	13.96	10.06	9.00	12.84
1982m05	16.13	13.69	10.02	9.10	12.67
1982m06	15.98	13.56	10.57	9.10	13.32
1982m07	15.96	13.20	10.46	9.30	12.97
1982m08	15.78	12.23	10.11	9.00	12.15
1982m09	15.76	11.40	9.91	8.70	11.48
1982m10	15.81	10.50	9.44	8.30	10.51
1982m11	15.79	10.64	8.28	8.10	10.18
1982m12	15.40	11.34	7.63	7.90	10.33
1983m01	14.99	11.60	7.39	7.60	10.37
1983m02	14.90	11.50	7.34	7.60	10.60
1983m03	14.65	11.02	7.50	7.40	10.34
1983m04	14.55	10.56	7.96	7.40	10.19
1983m05	14.55	10.65	8.33	7.60	10.21
1983m06	14.52	10.39	9.06	8.00	10.64
1983m07	14.04	10.95	8.70	8.10	11.10
1983m08	13.96	11.07	8.71	8.30	11.42
1983m09	14.03	10.67	8.64	8.30	11.26
1983m10	14.15	10.61	8.31	8.10	11.21
1983m11	14.12	10.29	8.39	8.10	11.32
1983m12	13.96	10.35	8.49	8.20	11.44

https://doi.org/10.26219/heal.aueb.5789

1984m01	13.76	10.28	8.45	8.20	11.29
1984m02	13.87	10.42	8.35	8.00	11.44
1984m03	13.98	10.23	8.08	7.90	11.90
1984m04	13.81	10.40	8.09	7.90	12.17
1984m05	13.82	10.93	8.32	8.00	12.89
1984m06	13.99	11.15	8.42	8.10	13.00
1984m07	13.91	11.67	8.60	8.10	12.82
1984m08	13.58	10.98	8.34	7.90	12.23
1984m09	13.05	10.78	8.09	7.60	11.97
1984m10	12.02	10.69	7.62	7.40	11.66
1984m11	12.36	10.32	7.45	7.20	11.25
1984m12	12.70	10.46	7.44	7.00	11.20
1985m01	12.12	10.96	7.38	7.10	11.15
1985m02	12.25	11.07	7.89	7.40	11.35
1985m03	12.27	10.90	8.02	7.60	11.78
1985m04	11.98	10.68	7.70	7.30	11.42
1985m05	11.72	10.88	7.54	7.10	10.96
1985m06	12.12	10.70	7.35	6.90	10.36
1985m07	11.87	10.44	7.24	6.70	10.51
1985m08	11.96	10.37	6.93	6.40	10.59
1985m09	11.97	9.97	6.89	6.30	10.67
1985m10	11.59	10.22	7.03	6.50	10.56
1985m11	11.22	10.37	6.98	6.60	10.08
1985m12	11.33	10.45	6.93	6.50	9.60
1986m01	10.79	10.80	6.80	6.30	9.51
1986m02	10.10	10.40	6.83	6.20	9.07
1986m03	9.29	9.39	6.46	5.90	8.13
1986m04	8.56	8.76	6.21	5.50	7.59
1986m05	8.58	9.00	6.42	5.80	8.02
1986m06	8.65	9.23	6.41	5.90	8.23
1986m07	8.55	9.37	6.22	5.90	7.86
1986m08	8.05	9.41	6.03	5.70	7.72
1986m09	8.51	9.97	6.03	5.80	8.08
1986m10	8.97	10.62	6.21	6.00	8.04
1986m11	9.49	10.80	6.29	6.10	7.81
1986m12	9.89	10.69	6.41	5.90	7.67
1987m01	9.54	10.09	6.21	5.80	7.60
1987m02	9.84	9.83	6.13	5.70	7.69
1987m03	9.42	9.16	6.21	5.60	7.62
1987m04	9.51	9.12	6.14	5.50	8.31
1987m05	9.86	8.82	6.11	5.40	8.79
1987m06	10.40	9.23	6.26	5.60	8.63

https://doi.org/10.26219/heal.aueb.5789

1987m07	10.39	9.23	6.40	5.80	8.70
1987m08	10.60	9.92	6.50	6.00	8.97
1987m09	11.06	9.98	6.72	6.20	9.58
1987m10	10.95	9.88	7.02	6.50	9.61
1987m11	10.46	9.20	6.37	6.00	8.99
1987m12	10.57	9.57	6.18	6.00	9.12
1988m01	9.73	9.57	6.06	6.00	8.82
1988m02	9.33	9.38	5.90	5.80	8.41
1988m03	9.61	9.12	5.80	5.70	8.61
1988m04	9.42	9.12	5.77	5.80	8.91
1988m05	9.38	9.27	5.99	6.10	9.24
1988m06	9.07	9.32	5.98	6.10	9.04
1988m07	9.18	9.51	6.30	6.40	9.20
1988m08	9.45	9.47	6.44	6.50	9.33
1988m09	9.00	9.60	6.23	6.30	9.06
1988m10	8.77	9.23	5.98	5.98	8.89
1988m11	8.87	9.30	5.99	5.99	9.07
1988m12	8.81	9.46	6.23	6.23	9.13
1989m01	8.77	9.35	6.44	6.50	9.07
1989m02	9.47	9.15	6.94	6.90	9.16
1989m03	9.21	9.26	7.08	6.90	9.33
1989m04	8.99	9.52	7.03	6.90	9.18
1989m05	9.00	9.52	7.34	7.10	8.86
1989m06	8.96	9.88	7.18	6.90	8.28
1989m07	8.80	9.53	7.05	6.80	8.19
1989m08	8.81	9.37	7.04	6.80	8.26
1989m09	9.18	9.62	7.27	7.10	8.31
1989m10	9.34	9.81	7.53	7.30	8.15
1989m11	9.53	9.99	7.85	7.60	8.03
1989m12	9.80	9.96	7.91	7.60	8.02
1990m01	10.12	10.28	8.36	7.90	8.39
1990m02	10.76	10.72	8.95	8.70	8.66
1990m03	10.21	11.46	9.13	8.90	8.74
1990m04	10.06	11.77	9.01	8.90	8.92
1990m05	10.05	11.49	9.05	8.90	8.90
1990m06	10.10	11.01	9.04	9.00	8.62
1990m07	10.04	11.03	8.80	8.70	8.64
1990m08	10.71	11.41	9.05	9.00	8.97
1990m09	11.11	11.32	9.21	9.10	9.11
1990m10	10.72	11.12	9.23	9.10	8.93
1990m11	10.58	10.94	9.17	9.00	8.60
1990m12	10.53	10.42	9.14	9.00	8.31

https://doi.org/10.26219/heal.aueb.5789

1991m01	10.10	10.22	9.23	9.10	8.33
1991m02	9.61	9.89	9.73	8.60	8.12
1991m03	9.66	10.06	8.73	8.50	8.38
1991m04	9.40	9.99	8.70	8.50	8.29
1991m05	9.30	10.15	8.73	8.50	8.33
1991m06	9.59	10.34	8.72	8.50	8.54
1991m07	9.66	10.10	8.90	8.80	8.50
1991m08	9.41	9.89	8.95	8.80	8.17
1991m09	9.30	9.54	8.86	8.70	7.96
1991m10	9.26	9.62	8.82	8.60	7.88
1991m11	9.31	9.68	8.85	8.60	7.83
1991m12	9.19	9.56	8.83	8.50	7.58
1992m01	8.87	9.34	8.48	8.10	7.48
1992m02	8.86	9.21	8.38	8.10	7.78
1992m03	9.13	9.54	8.43	8.10	7.93
1992m04	9.08	9.33	8.47	8.20	7.88
1992m05	8.92	8.99	8.46	8.30	7.80
1992m06	9.15	9.02	8.42	8.30	7.72
1992m07	9.52	8.88	8.45	8.30	7.40
1992m08	9.57	9.13	8.50	8.40	7.19
1992m09	8.98	9.16	8.18	8.00	7.08
1992m10	8.63	9.24	7.62	7.40	7.26
1992m11	8.62	8.83	7.44	7.30	7.43
1992m12	8.14	8.84	7.35	7.30	7.30
1993m01	8.32	8.91	7.07	7.00	7.17
1993m02	8.05	8.63	6.85	6.80	6.89
1993m03	7.75	8.33	6.35	6.40	6.65
1993m04	7.58	8.39	6.59	6.50	6.64
1993m05	7.48	8.60	6.63	6.60	6.68
1993m06	7.02	8.39	6.61	6.60	6.55
1993m07	7.01	7.96	6.42	6.40	6.34
1993m08	6.50	7.38	6.19	6.20	6.18
1993m09	6.45	7.18	6.00	6.00	5.94
1993m10	6.23	7.09	5.81	5.80	5.90
1993m11	6.25	7.06	5.68	5.60	6.25
1993m12	5.86	6.46	5.56	5.50	6.27
1994m01	5.96	6.41	5.43	5.50	6.24
1994m02	6.43	6.83	5.67	5.80	6.44
1994m03	6.76	7.47	6.15	6.20	6.90
1994m04	7.06	7.83	6.35	6.30	7.32
1994m05	7.25	8.23	6.53	6.40	7.47
1994m06	7.87	8.55	6.98	6.90	7.43

https://doi.org/10.26219/heal.aueb.5789

1994m07	7.49	8.41	6.81	6.70	7.61
1994m08	8.04	8.52	7.02	6.90	7.55
1994m09	8.33	8.72	7.37	7.30	7.81
1994m10	8.50	8.63	7.43	7.40	8.02
1994m11	8.14	8.53	7.40	7.30	8.16
1994m12	8.37	8.44	7.44	7.40	7.97
1995m01	8.22	8.61	7.53	7.50	7.93
1995m02	8.15	8.52	7.35	7.30	7.69
1995m03	7.99	8.50	7.14	7.10	7.52
1995m04	7.96	8.39	6.87	6.80	7.41
1995m05	7.58	8.18	6.63	6.50	6.99
1995m06	7.77	8.16	6.54	6.40	6.59
1995m07	7.49	8.36	6.60	6.50	6.71
1995m08	7.40	8.24	6.47	6.40	6.90
1995m09	7.65	8.09	6.25	6.10	6.63
1995m10	7.61	8.29	6.22	6.10	6.43
1995m11	7.19	7.97	5.94	5.80	6.31
1995m12	6.89	7.75	5.70	5.60	6.11
1996m01	6.77	7.79	5.45	5.30	6.07
1996m02	7.00	8.10	5.82	5.70	6.28
1996m03	6.92	8.34	6.04	5.90	6.72
1996m04	6.76	8.30	5.88	5.80	6.94
1996m05	6.71	8.35	5.87	5.80	7.08
1996m06	6.84	8.36	6.03	6.00	7.20
1996m07	6.59	8.25	6.04	6.00	7.13
1996m08	6.62	8.16	5.85	5.70	6.94
1996m09	6.20	8.16	5.67	5.60	7.13
1996m10	6.11	7.87	5.41	5.40	6.87
1996m11	5.79	7.81	5.36	5.30	6.55
1996m12	5.82	7.71	5.28	5.20	6.63
1997m01	5.69	7.74	5.21	5.10	6.89
1997m02	5.39	7.39	5.01	4.90	6.76
1997m03	5.80	7.62	5.23	5.10	7.03
1997m04	5.93	7.77	5.34	5.20	7.18
1997m05	5.96	7.23	5.26	5.10	7.00
1997m06	5.67	7.18	5.20	5.00	6.82
1997m07	5.50	6.97	5.10	4.90	6.55
1997m08	5.65	7.02	5.27	5.10	6.64
1997m09	5.55	6.79	5.24	5.10	6.54
1997m10	5.80	6.48	5.32	5.30	6.37
1997m11	5.66	6.54	5.33	5.30	6.18
1997m12	5.45	6.31	5.12	5.10	6.06

https://doi.org/10.26219/heal.aueb.5789

1998m01	5.26	6.07	4.88	4.80	5.87
1998m02	5.11	5.99	4.76	4.70	5.94
1998m03	5.04	5.87	4.76	4.70	6.00
1998m04	5.12	5.70	4.84	4.70	5.98
1998m05	5.05	5.75	4.89	4.80	5.99
1998m06	4.95	5.54	4.75	4.60	5.78
1998m07	4.91	5.56	4.66	4.50	5.76
1998m08	4.61	5.46	4.44	4.30	5.64
1998m09	4.39	5.06	4.16	4.00	5.34
1998m10	4.51	4.98	4.15	4.00	5.24
1998m11	4.43	4.85	4.17	4.00	5.43
1998m12	4.41	4.54	3.91	3.70	5.29

Table 2 (Order-Aic)

Order of VAR	AIC
1	45.57
2	0.00
3	15.3
4	21.85
5	32.87
6	67.98
7	91.52
8	102.44
9	123.68
10	182.39
11	195.45

https://doi.org/10.26219/heal.aueb.5789

https://doi.org/10.26219/heal.aueb.5789

REFERENCES

Amos, D.E. & Koopmans, L.H. (1963) *Tables of distribution of coherences for stationary bivariate Gaussian processes*, Sandia Corporation Monograph

Brockwell, Peter J & Davis Richard, A. (1991) *Time Series: Theory and Methods.* Springer-Verlag,

Bartlett. (1955) *An Introduction to Stochastic Processes,* Cambridge University Press

Dickey, D.A. & W.A. Fuller (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of American Statistical Association,* 74, 427-431

Engle, R.F. & C.W.J. Granger (1986) Co-integration and Error Correction: Representation, Estimation and Testing, *Econometrica,* 55, 251-276

Hamilton, James D. (1994), *Time Series Analysis,* Princeton. Univ. Press.

Hannan, E. J. (1970) *Multiple Time Series and Methods,* John Wiley, New York.

Hansen, B.E. (1992) Efficient Estimation and Testing of Cointegrating Vectors in the Presence of Deterministic Trends, *Journal of Econometrics,* 53, 87-121

Ioannidis, E. E. Spectra of Bivariate VAR (p) Models, 2007, Journal of Statistical Planning and Inference,137, 554-566.

Johansen, S. (1988) Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control,* 12, 231-254

Johansen, S. (1995) *Likelihood Based Inference in Cointegrated Vector Error Correction Models, Oxford University Press*, Oxford

Juselius, Katarina (2001). Big Shocks, Outliers, and Interventions. A Cointegration and Common Trends Analysis of Daily Bond Rates.

Mac Kinnon, James G. (1996) Numerical Distribution Functions for Unit Root and Cointegration Tests, *Journal of Applied Econometrics***,** 11, 601-618

Maddale, S. (1999) *Unit Root and Cointegration.* Cambridge Univ. Press

Osterwald – Lenum, M. (1992) *A Note with Quantiles of Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Statistics*, Oxford Bulletin of Economics and Statistics

Parzen, E (1974) *Some recent advances in time series modelling*, IEEE Transactions on Automatic Control

Phillips, P.C.B. (1986) Understanding Spurious Regression in Econometrics, *Journal of Econometrics,* 33, 311-340

Phillips, P.C.B. & P. Perron (1988) Testing for a Unit Root in Time Series Regression, *Biometrika,* 75, 345-346

Phillips, P.C.B. & S. Quliaris (1990) Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica,* 58, 73-93

Phillips, P.C.B. (1991) Optimal Inference in Cointegrated Systems, *Econometrica,* 59, 283-306

Priestley, M. B. (1983) *Spectral Analysis and Time Series,* Academic Press, New York

Reinsel, G.C. & S.K. Ahn (1992) Vector Autoregression Models with Unit Rots and Reduced Rank Stracture: Estimation, Likelihood Ratio Test and Forecasting, *Journal of Time Series Analysis,* 13, 353-375

Reimars H.-E. (1992) Comparisons of Tests for Multivariate Cointegration, *Statistical Papers,* 33, 335-359

Tukey, J. (1949) *The sampling theory of power spectrum estimates, Proc. Symp. On Applications of Autocorrelation Analysis to Physical Problems*, NAVEXOS, Office of Naval Research, Washington

Zivot, Eric and Wang, Jiahui. (2006) *Modeling Financial Time Series with S-Plus,* Springer