Università
Ca'Foscari
Venezia

Master's Degree programme – Second Cycle
(*D.M. 270/2004*)
in Economics and Finance

Final Thesis

—

Ca' Foscari
Dorsoduro 3246
30123 Venezia

Macroeconomic Factors and the U.S. Stock Market Index: A Cointegration Analysis

Supervisor
Ch. Prof. Stefano Federico Tonellato

Graduand
Verena Brufatto
Matriculation Number 849145

Academic Year
2014 / 2015

Abstract

This thesis is aimed at investigating the impact that changes and shocks in a set of selected macroeconomic variables have on the U.S. stock market returns. The existence of a long-run equilibrium relationship between fundamentals and the stock market index is inquired using the methodological framework of cointegration analysis and a vector error correction model. Moreover, the short-run dynamic dependencies between the variables are examined by performing an impulse response analysis and the empirical question of whether economic variables are useful indicators of future stock market returns is addressed. The empirical results of this study indicate that the U.S. stock market index establishes a cointegrating relationship with some of the selected macroeconomic variables, showing that information about relevant economic indicators is reflected in stock returns and that changes in fundamentals are significantly priced in the stock market index. In addition, the impulse response analysis highlights the presence of meaningful short-run dynamic effects, on the grounds that innovations in the macroeconomic variables are seen to exert an impact on stock prices.

Contents

Abbreviations

Table 1: List of abbreviations

Abbreviation	Definition
ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller Unit Root Test
AIC	Akaike Information Criterion
APT	Arbitrage Pricing Theory
AR (AR(p))	Autoregressive Process (of order p)
ARCH	Autoregressive Conditional Heteroskedasticity (Model)
ARMA (ARMA($p,\ q$))	Autoregressive Moving Average Process (of order $p,\ q$)
BIC	Bayesian-Schwartz Information Criterion
CRDW	Cointegration Regression Durbin-Watson Test
DF	Dickey-Fuller Unit Root test
DOLS	Dynamic Ordinary Least Squares (Estimator)
DGP	Data Generating Process
DW	Durbin-Watson Statistic
EC / ECM	Error Correction / Error Correction Model
ECT	Error Correction Term
EG	Engle-Granger Cointegration Test
FEVD	Forecast Error Variance Decomposition
GARCH	Generalized Autoregressive Conditional Heteroskedasticity Model
GDP	Gross Domestic Product
GNP	Gross National Product

HQIC	Hannan-Quinn Information Criterion
IID	Independently and Identically Distributed
IRF	Impulse Response Function
KPSS	Kwiatkowski-Phillips-Schmidt-Shin Stationarity Test
LR	Likelihood Ratio Test
LM	Lagrange Multiplier Test
MA (MA(q))	Moving Average Process (of order q)
ML / MLE	Maximum Likelihood / Maximum Likelihood Estimator
MSE	Mean-Squared Errors
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary Least Squares (Estimator)
PP	Phillips-Perron Unit Root Test
PACF	Partial Autocorrelation Function
RHS / LHS	Right Hand Side / Left Hand Side
S&P500	Standard and Poor's 500 Stock Index
VAR (VAR(p))	Vector Autoregressive Model (of order p)
VEC / VECM	Vector Error Correction / Vector Error Correction Model

Introduction

The empirical question of whether fundamentals can be influential factors in the determination and prediction of stock prices is, by now, well-documented in the literature. Although the establishment of a link between macroeconomic variables and stock prices is in contrast with the efficient market hypothesis by Fama (1970)[1], evidence of a long-run relationship between fundamentals and stock markets across a variety of countries was found by a number of studies. As maintained by economic theory and previous surveys, macroeconomic forces are believed to drive stock prices on the grounds that changes in economic activities are expected to affect corporate future earnings, future dividend payouts and discount rates, which in turn influence the equity market (Fama (1981, 1990), Chen et al. (1986), Ferson and Harvey (1991, 1993)). Much of the early literature in this thematic area is characterized by the use of the arbitrage pricing theory (APT) model by Ross (1976) for investigating whether a series of risk factors can explain stock returns[2]. Employing this methodology, evidence that macroeconomic variables have a systematic effect on U.S. stock market returns was found by Fama and Schwert (1977), Nelson (1977) and Jaffe and Mandelker (1976), among others. A study by Chen et al. (1986), based on the APT model, reports that the set of macroeconomic variables composed of interest rates spread, expected and unexpected inflation, industrial production and bond yield spread signif-

[1]The efficient market hypothesis states that, if the market is efficient, the possibility of earning abnormal profits through the prediction of future movements of stock prices is precluded to investors, since, due to competition among investors, any relevant information enclosed in changes in fundamentals will be swiftly incorporated in current stock prices.

[2]The APT approach focuses on establishing whether changes in fundamental variables represent significant risk factors, rewarded by risk premia, on the equity market and if they are consistently priced in stock market returns.

icantly influences stock market returns. The findings of Chen et al. suggested the possibility that a long-term equilibrium relationship between fundamentals and stock prices existed, but it was not until the development of cointegration analysis by Granger (1986) and Johansen and Juselius (1990) that the basis for the empirical determination of this relationship was set. In particular, the introduction of the error correction model (ECM) by Engle and Granger (1987) and of the vector error correction model (VECM) by Johansen (1991) offered the opportunity to investigate the dynamic co-movements between the variables and to examine the adjustment mechanism to previous periods departures from long-run equilibrium, even in the presence of non-stationary time series. Ever since its development, the methodology of cointegration analysis has been widely applied in economic studies with focus on the stock exchanges of industrialized countries as well as emerging economies. Several studies concentrated on examining the co-movements of macroeconomic variables and the stock exchange in the U.S. and Canada (Fama (1981), Geske and Roll (1983), Chen et al. (1986), Darrat (1990), Lee (1992), Abdullah and Hayworth (1993), McMillan (2001)), others tried to establish a link between real activity and the U.K. equity market (Poon and Taylor (1991), Cheng (1995)), while some studies focused on finding this relationship in European countries (Wasserfallen (1989), Asprem (1989), Gjerde and Saettem (1999), Nasseh and Strauss (2000), Hondroyiannis and Papapetrou (2001), Panetta (2002), Tsoukalas (2003), Dritsaki and Adamopoulos (2005)), in Asian countries (Hamao (1988), Mukherjee and Naka (1995), Kwon and Shin (1999), Ibrahim (1999), Maysami and Koh (2000), Maysami and Sims (2001a, 2001b, 2002), Maysami et al. (2004)), in India (Padhan (2007), Agrawalla and Tuteja (2007), Naik and Phadi (2012)) and in African countries (Jefferis and Okeahalam (2000), Lekobane-Sikalao and Lekobane (2012), Ogbulu et al. (2014)).

Although the results of the aforementioned surveys indicate the presence of a strong link between the stock market performance and fundamentals, there is a certain number of studies that disclaim the existence of any significant relationship. Culter et al. (1988), for example, debate the hypothesis that information about expected corporate cash flows and discount rates are perfectly incorporated into stock prices and find that the influence of macroeconomic news on stock markets is generally weak. The effort of Schwert (1989) to relate stock volatility in the U.S. to changes in leading economic indicators failed to find strong evidence that economic factors contribute to the determination of highly volatile future stock prices. A survey by Richards (1996) reports that the predictability of stock returns in emerging equity markets appears to have diminished over time, while Allen and Jagtianty (1997) argue that, due to the widespread practice of writing derivative contracts on interest rates for hedging purposes, the sensitivity of the stock market to changes in interest rates suffered a decrease between the late 1980's and the early 1990's. According to Pearce (1983), the stock market is likely to generate false signals about the direction of the economy – as it was the case when, during the 1987 stock market crash, the advent of an economic recession was wrongly prognosticated by the market – and should hence not be regarded as a reliable economic indicator. Nevertheless, evidence that macroeconomic factors significantly contribute to the determination of stock prices was found by several cointegration-based studies. For example, Cheung and Ng (1998) investigated the long-run relationship between five national stock market indexes and aggregate real economic variables such as real GNP, real oil price, real money supply and real consumption. Using Johansen's likelihood ratio test for the cointegration rank and an ECM, the authors established that the stock market indexes of Canada, Germany, Italy, Japan and the U.S. are strongly related to changes in real domestic aggregate activity. By employing Johansen's VECM, Mukherjee and Naka (1995) report that the Japanese stock market is cointegrated with a set of domestic macroeconomic variables –that is, industrial production, money supply, exchange rate, call money rate, long-term interest rate and inflation. In a series of papers, Maysami and Sims (2001a, 2001b, 2002) examined the relationship between money supply, interest rates, inflation, exchange rate, real activity and the stock exchanges of Malaysia and Thailand, Japan and Korea and Hong Kong and Singapore using a vector error correction model (VECM). In order to proxy for the effect of the 1997 Asian financial crisis, a dummy variable was included in the set of macroeconomic factors. The output of the studies highlights the presence of a significant long-run relationship between fundamentals and stock markets, which however differs in type and size depending on the particular financial structure of the country under analysis. Through Johansen's cointegration rank test and a forecast error variance decomposition, Nasseh and Strauss (2000) found that industrial production, short- and long-term interest rates, business surveys of manufacturing and foreign stock prices strongly influenced the stock market prices of six European countries, namely Germany, Italy, France, Netherlands, the U.K. and Switzerland. McMillan (2001) inquired the existence of a cointegrating relationship between two U.S. stock market indexes and industrial production, inflation, short- and long-term interest rates and money supply. The results indicated a significant long-run relationship between the Standard and Poor's

500 (S&P500) and the Dow Jones Industrial Average (DJX) and industrial production, long-term interest rates and inflation. A variance decomposition highlighted the contribution of industrial production, short- and long- term interest rates and inflation to explaining stock price variability in the U.S. With an application to the Australian stock exchange, Chaudhuri and Smile (2004) focused on determining the effect of changes in real macroeconomic activity on real stock returns. Using the methodology of cointegration analysis, the authors found evidence that real domestic aggregate activity along with stock return variability in New Zealand and the U.S. have significant explanatory power over stock market movements in Australia.

The purpose of this survey is to contribute to empirically assess the extent to which a long-run relationship between U.S. stock prices and a set of macroeconomic variables exists. In particular, the question of whether current economic activities are among the determinants of U.S. stock market prices and whether they are influential factors in predicting future stock returns is of interest. If stock market returns consistently reflect macroeconomic information, the U.S. stock market index should be cointegrated with the set of macroeconomic variables and changes in the latter should contribute significantly to the cointegrating relationship. In economic terms, this would imply that the U.S. equity market is sensitive to changes in fundamentals and that future stock prices can be determined, to some degree, by changes in economic factors. The methodological framework of cointegration analysis and a vector error correction model are used in an attempt to capture the long-run response of the U.S. stock market, represented by the Dow Jones Industrial Average, to changes in the levels of fundamental economic variables such as industrial production, money supply, short-term interest rates, crude oil price and the rate of inflation. This study finds that the selected macroeconomic variable and the U.S. stock index are cointegrated, indicating that the equity market incorporates economic information into the stock price index. The existence of a vector error correction model implies that there is Granger causality from macroeconomic factors to the stock market index and consequently, past values of the set of macro-variables possess a certain forecasting ability with respect to U.S. stock prices. By performing an impulse-response analysis, the short-term effect of a shock in each of the macroeconomic factors on the stock market is investigated. The results of this survey highlight the presence of significant short-term relations between the U.S. stock index and the set of macroeconomic factors, in addition to a long-term equilibrium relationship.

Chapter 1

Time Series Concepts

Univariate time series models, which will be briefly introduced later on in this section, are commonly used to model and forecast economic and financial variables by relying on the information retained in the variable's own past values and current or past values of an error process. This class of models represents and alternative to structural models, which are typically multivariate and which are aimed at explaining how the value of a variable changes in response to the movements of a set of explanatory variables.

1.1 Stationarity, Ergodicity and Weak Dependence

The following section is intended to outline some of the fundamental concepts in time series analysis, prior to introducing the main notions of this survey in the next chapters.

Definition 1 (Stochastic Process). *A sequence $\{Y_t(\omega)\}_{t\in\mathbb{T}}$ of random variables or random vectors, which are all defined on the same probability space $(\Omega, \mathscr{A}, \mathbb{P})$, is called a stochastic process. The time parameter $\mathbb{T}$ usually corresponds to the set of natural, real or integer numbers.*

A stochastic process depends both on time and on uncertainty. A realization $\{y_t\}_{t\in\mathbb{T}}$ of a random variable is the observed value of one of the variables that constitute the stochastic process. A sequence of realizations is what is called a *time series*. Since normally the process is not observed at every point in time in $\mathbb{T}$, but only on a finite subset T of $\mathbb{T}$, the time series, i.e. the observed path, has length T. Often the terms stochastic process and time series are used as synonyms to indicate a series of random variables which are indexed by time; the notation $\{y_t\} = \{...y_{t-1}, y_t, y_{t+1}...\}$ for $t = 1, ..., T$ can be used equivalently to $\{Y_t(\omega)\}_{t\in\mathbb{T}}$ to denote a stochastic process. Time series data are very frequently used in economic empirical studies. The concept of stationarity plays a crucial role in the analysis of financial time series. When considering cross sectional data, i.e. data which stems from the same point in time, the samples are $i.i.d.$, whereas this is not necessarily true for sample variables from a time series. As a consequence, estimators like the sample mean might no longer be unbiased and consistent when applied to time series data. In order to employ usual estimation methods on a sample from a time series, the stochastic process is required to be stationary and ergodic. Stationarity provides the condition for the series to be identically distributed, while ergodicity ensures that certain results, such as the law of large numbers and the central limit theorem, hold even when the realizations are not independent.

Definition 2 (Strict Stationarity)**.** *Considering the points in time $t_1, ..., t_n \in \mathbb{T}$, for $n \in \mathbb{N}$, the stochastic process $\{Y_t(\omega)\}_{t\in\mathbb{T}}$ is strictly stationary if, for $h \in \mathbb{T}$:*

$$P(Y_{t_1} \leq y_1, ..., Y_{t_n} \leq y_n) = P(Y_{t_1+h} \leq y_1, ..., Y_{t_n+h} \leq y_n)$$

A process is defined strictly stationary when its probability distribution is stable over time, in the sense that the joint distribution of a collection of random variables from the process remains unchanged when the collection is shifted in time. The stochastic process $\{Y_t(\omega)\}_{t\in\mathbb{T}}$, which from now on we will denote $\{y_t\}$ for simplicity, is strictly stationary if, for all h, the joint distribution of $(y_{t_1}, ..., y_{t_n})$ is the same as the joint distribution of $(y_{t_1+h}, ..., y_{t_n+h})$. In other words, y_t has the same probability distribution as y_{t+h}, which is the value of the process h periods of time ahead. The same holds for any pair or sequence of terms in the process, up to an arbitrary positive integer n. Usually, the assumption of a weaker form of stationarity, which is more manageable than strict stationarity, suffices for analysis purposes.

Definition 3 (Weak Stationarity)**.** *A stochastic process $\{y_t\}$ with a finite second moment $E[y_t^2 < \infty]$ is covariance stationary if $E[y_t] = \mu_t = \mu$ and if $\mathrm{Cov}[y_t, y_{t-h}] = E[(y_t - \mu)(y_{t-h} - \mu)] = \gamma_h$, for all $t \in \mathbb{T}$ and for any h.*

A process is weakly stationary when its moments are independent of time. This definition implies that also the variance is invariant under a time shift, since $\mathrm{Var}[y_t] = \mathrm{Cov}[y_t, y_t] = \gamma_0$.

Hence, the first two moments of a covariance stationary process are constant — since they don't depend on t, while the covariance, and consequently also the correlation between y_t and y_{t-h}, only depends on the time gap h and not on the starting point t. The autocovariance of a process $\{y_t\}$, denoted as

$$\mathrm{Cov}[y_t, y_{t-h}] = \gamma_h = E[(y_t - \mu)(y_{t-h} - \mu)] \tag{1.1}$$

is, in the case of a weakly stationary process, the covariance between the value of y at present time t and its value at previous times. The correlation between current and past values of the process Y_t is known as the autocorrelation function of Y_t and is obtained by divinding the autocovariance function γ_h by the variance γ_0. In particular, the correlation coefficient between Y_t and Y_{t-h} is defined as the lag-h autocorrelation, written

$$\rho_h = \frac{\mathrm{Cov}[y_t, y_{t-h}]}{\sqrt{\mathrm{Var}[y_t]\mathrm{Var}[y_{t-h}]}} = \frac{\mathrm{Cov}[y_t, y_{t-h}]}{\mathrm{Var}[y_t]} = \frac{\gamma_h}{\gamma_0} \tag{1.2}$$

where $\mathrm{Var}[y_t] = \mathrm{Var}[y_{t-h}]$, since y_t is a weakly stationary process. By plotting the autocorrelations ρ_h against the time lags $h = 1, 2, ...$, the autocorrelation function (ACF), also known as correlogram, is built. Gaussian white noise processes are known to be stationary processes. A time series $\{y_t\}$ with finite first and second moment is said to be white noise, while if $\{y_t\}$ follows a normal distribution with mean 0 and variance σ^2, then it is Gaussian white noise and its realizations are i.i.d. Since the autocovariance of white noise processes is always zero except for at lag zero, where it is equal to the variance, observations are uncorrelated across time. Consequently, the ACF of a white noise sequence is 1 for $h = 0$ and zero otherwise. A white noise series does not display any kind of trending behavior, such that, in a time plot, the process will frequently cross it mean value of zero.

Definition 4 (Ergodicity)**.** *Let $(y_1, .., y_T)$ be a collection of T random variables from a weakly stationary stochastic process $\{y_t\}$ with expected value μ. The process is said to be mean-ergodic or first-order ergodic if*

$$p \lim \left(\frac{1}{T} \sum_{t=1}^{T} y_t \right) = \mu$$

A stationary stochastic process is first-order ergodic when the sample mean $\frac{1}{T}\sum_{t=1}^{T} y_t$ converges in probability to the population mean μ. If the same holds for both the first and second sample moments, the process is said to be ergodic in the wide sense and consistent estimation of population moments is ensured when the number of points in time at which the process is observed increases. Hence, when considering a long enough sample string, the time average of its elements is consistent for the population mean, provided that the realizations of the process are not too strongly correlated. Ergodicity also ensures that the law of large numbers can be applied to a sequence of dependent random variables stemming from a time series.

Definition 5 (Weak Dependence)**.** *A covariance stationary time series $\{y_t\}$ is said to be weakly dependent, if* $\mathrm{Corr}[y_t,\ y_{t+h}] \to 0$ *as* $h \to \infty$.

The concept of weak dependence is used to determine how strongly two realizations of a stochastic process are correlated when they are set far apart in time. For a stationary time series, weak dependence is given when y_t and y_{t+h} are almost independent for a large h, which implies that the lag-h autocorrelation of y_t decays sufficiently rapidly to zero as the number of lags h increases. Intuitively, any sequence that is i.i.d. is also weakly dependent. Because the autocorrelation of a covariance stationary process doesn't depend on the starting point t, such processes are automatically weakly dependent. Weakly dependent time series are also said to be asymptotically uncorrelated. In practice, as the time distance between two random variables increases, their correlation progressively diminishes, until it eventually goes to zero when the number of lags reaches infinity. Weak dependence has useful implications in regression analysis, for it ensures the validity of the law of large numbers and of the central limit theorem, replacing the assumption of random sampling in time series data.

1.2 Stationary Time Series Models

1.2.1 Autoregressive Processes

AR(1) Models

An autoregressive model of order one, or AR(1) in short, has the form

$$y_t = \phi_0 + \phi_1 y_{t-1} + u_t \tag{1.3}$$

where the term u_t is a zero mean white noise process with variance σ^2. As can be seen from the model expression, in an AR(1) model the value of y_t depends on its first lag y_{t-1} and on the disturbance term u_t. If we assume that (1.3) is a covariance stationary series, then it must be that $E[y_t] = \mu$, $\mathrm{Var}[y_t] = \gamma_0$ and $\mathrm{Cov}[y_t, y_{t-h}] = \gamma_h$, where μ and γ_0 are constants and γ_h only depends on *h* and not on *t*. The expected value of (1.3) is calculated as

$$E[y_t] = \mu = \frac{\phi_0}{1 - \phi_1} \tag{1.4}$$

since, under the weak stationarity assumption, $E[y_t] = E[y_{t-1}] = \mu$, from which it follows that $\mu = \phi_0 + \phi_1\mu$ or equivalently $\phi_0 = (1 - \phi_1)\mu$. From (1.4) it can be seen that the mean of y_t only exists as long as $\phi_1 \neq 1$ and is non-zero only for $\phi_0 \neq 0$. By the same logic, $\mathrm{Var}[y_t] = \mathrm{Var}[y_{t-1}]$ and the variance of y_t is

$$\mathrm{Var}[y_t] = \gamma_0 = \frac{\sigma^2}{1 - \phi_1^2}$$

since y_{t-1} and u_t are uncorrelated. The variance of the AR(1) process is positive and bounded provided that $\phi_1^2 < 1$, such that the weak stationarity of (1.3) is ensured by the constraint $|\phi_1| < 1$. Exploiting the fact that $\phi_0 = (1-\phi_1)\mu$, the lag-*h* autocovariance of y_t is calculated by rewriting (1.3) as

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + u_t \tag{1.5}$$

and then multiplying each side of the previous equation by $(y_{t-h} - \mu)$ and taking expectations, yielding

$$E[(y_t - \mu)(y_{t-h} - \mu)] = \phi_1 E[(y_{t-1} - \mu)(y_{t-h} - \mu)] + E[u_t(y_{t-h} - \mu)]$$

Since $E[u_t(y_t - \mu)] = \sigma^2$ and $E[u_t(y_{t-h} - \mu)] = 0$, the lag-$h$ autocovariance of y_t is

$$\gamma_h = \begin{cases} \phi_1\gamma_1 + \sigma^2 & \text{if } h = 0 \\ \phi_1\gamma_{h-1} = \frac{\sigma^2\phi_1^h}{1-\phi_1^2} & \text{if } h > 0 \end{cases}$$

Considering that $\rho_h = \gamma_h/\gamma_0$, the autocorrelations of the AR(1) model are equal to $\rho_0 = 1$, $\rho_1 = \phi_1$, $\rho_2 = \phi_1^2$,..., $\rho_h = \phi_1^h$. Equivalently, the lag-h autocorrelation of y_t can be expressed as

$$\rho_h = \phi_1\rho_{h-1} \tag{1.6}$$

such that it satisfies

$$(1 - \phi_1 L)\rho_h = 0 \tag{1.7}$$

where L is the lag operator and $L\rho_h = \rho_{h-1}$. The first-order polynomial equation

$$1 - \phi_1 z = 0 \tag{1.8}$$

is known as the characteristic equation of the AR(1) model. The root of the characteristic equation is referred to as the characteristic root of the model and it is calculated as the inverse of the solution of (1.8). The stationarity of y_t is ensured when the characteristic root lies inside the complex unit circle or equivalently, when its absolute value is less than one. By solving (1.8) with respect to z and then taking the inverse, the root $\lambda = 1/z = \phi_1$ is obtained, so that (1.3) is a stationary process provided that $|\phi_1| < 1$. In this instance, the ACF of y_t will decay exponentially to zero as h gets large. For $\phi_1 > 0$, the convergence of the correlogram will be direct, while the ACF will waver around zero for $\phi_1 < 0$.

AR(2) Models

The following process

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t \tag{1.9}$$

is called an autoregressive model of order two, written AR(2). Assuming that the process is weakly stationary, the expected value of (1.9) is calculated as

$$E[y_t] = \mu = \frac{\phi_0}{1 - \phi_1 - \phi_2}$$

which exists as long as $\phi_1 + \phi_2 \neq 1$. Exploiting the fact that $\phi_0 = \mu(1 - \phi_1 - \phi_2)$, the AR(2) model can be rewritten as

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + u_t \tag{1.10}$$

The higher order moments of the AR(2) model are obtained by multiplying (1.10) by $(y_{t-h} - \mu)$ for $h = 0, 1, 2, ...$, and taking expectations[1]

$$\begin{aligned}
E[(y_t - \mu)(y_t - \mu)] &= \phi_1 E[(y_{t-1} - \mu)(y_t - \mu)] + \phi_2 E[(y_{t-2} - \mu)(y_t - \mu)] + E[u_t(y_t - \mu)] \\
E[(y_t - \mu)(y_{t-1} - \mu)] &= \phi_1 E[(y_{t-1} - \mu)(y_{t-1} - \mu)] + \phi_2 E[(y_{t-2} - \mu)(y_{t-1} - \mu)] + E[u_t(y_{t-1} - \mu)] \\
&\vdots \\
E[(y_t - \mu)(y_{t-h} - \mu)] &= \phi_1 E[(y_{t-1} - \mu)(y_{t-h} - \mu)] + \phi_2 E[(y_{t-h} - \mu)(y_{t-1} - \mu)] + E[u_t(y_{t-h} - \mu)]
\end{aligned} \tag{1.11}$$

which yields

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2 \tag{1.12}$$

$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 \tag{1.13}$$

$$\gamma_h = \phi_1\gamma_{h-1} + \phi_2\gamma_{h-2} \tag{1.14}$$

By dividing (1.13) and (1.14) by γ_0, it is possible to obtain the autocorrelations of the AR(2) model

$$\rho_1 = \phi_1\rho_0 + \phi_2\rho_1 = \phi_1 + \phi_2\rho_1 \tag{1.15}$$

$$\rho_h = \phi_1\rho_{h-1} + \phi_2\rho_{h-2} \tag{1.16}$$

since $\rho_0 = 1$. By looking at equation (1.16), it can be seen that the autocorrelations satisfy $(1 - \phi_1 L - \phi_2 L^2)\rho_h = 0$, such that the second-order polynomial equation

$$1 - \phi_1 z - \phi_2 z^2 = 0 \tag{1.17}$$

represents the characteristic equation of y_t. As in the previous case of a AR(1) model, the two characteristic roots λ_1 and λ_2 of the AR(2) must be less than one in modulus in order for y_t to be stationary. Under this condition, (1.16) ensures that the ACF of the AR(2) will converge to zero

[1] This procedure is known as the Yule-Walker equations

as h gets large. The solutions of (1.17) are obtained by solving

$$z = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}$$

and the inverses of these solutions represent the characteristic roots. In case that both the characteristic roots are real numbers, the polynomial $(1 - \phi_1 L - \phi_2 L^2)$ can be factored as $(1 - \lambda_1 L)(1 - \lambda_2 L)$ and the ACF of the AR(2) model will be a mixture of two exponential decays. If instead λ_1 and λ_2 are complex numbers, the correlogram will display a sinusoidal path.

AR(p) Models

Let us generalize the results obtained for the AR(1) and AR(2) models by including p lags of the variable y_t in the model specification. The resulting process

$$\begin{aligned} y_t &= \phi_0 + \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + u_t \\ &= \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + u_t \end{aligned} \tag{1.18}$$

is known as an autoregressive model of order p, written AR(p)[2]. Using the lag operator and setting $\phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$, (1.18) can be rewritten as

$$\phi(L) y_t = \phi_0 + u_t \tag{1.19}$$

The AR(p) model is stationary when the solutions of the associated characteristic equation

$$1 - \phi_1 z - ... - \phi_p z^p = 0 \tag{1.20}$$

have modulus greater than one, or otherwise when the characteristic roots —that is, the inverses of the solutions of (1.20), are less than one in modulus. If stationarity is ensured, the ACF of the AR(p) model will decay exponentially to zero as the number of lags increases. Depending on the nature of the characteristic roots of y_t, the correlogram will exhibit a pattern of exponential decays and sinusoidal behavior. The expected value of an AR(p) model is equal to

$$E[y_t] = \frac{\phi_0}{1 - \phi_1 - ... - \phi_p} \tag{1.21}$$

while its other moments can be calculated as in the AR(2) case by means of the Yule-Walker equations

$$\gamma_0 = \phi_1 \gamma_1 + ... + \phi_p \gamma_p + \sigma^2 \tag{1.22}$$

[2]Chapter 2 treats the order selection, parameter estimation and model adequacy tests for AR models.

$$\gamma_h = \phi_1 \gamma_{h-1} + ... + \phi_p \gamma_{h-p} \tag{1.23}$$

$$\rho_h = \phi_1 \rho_{h-1} + ... + \phi_p \rho_{h-p} \tag{1.24}$$

According to the Wold decomposition theorem, any weakly stationary process can be represented in the form of an infinite order moving average process MA(∞). In the case of a zero mean AR(p) process that contains no constant term

$$\phi(L) y_t = u_t \tag{1.25}$$

Wold's decomposition is stated as

$$y_t = \psi(L) u_t \tag{1.26}$$

where $\psi(L) = \phi(L)^{-1} = (1 - \phi_1 L - ... - \phi_p L^p)^{-1}$.

1.2.2 Moving Average Processes

As we have seen in the previous section, in an AR model the value of the process y at present time t depends on its own previous values plus an error term. In contrast, in a moving average (MA) model, y_t can be seen as a weighted average of present and past disturbance terms. A MA model of order one (MA(1)) has the following expression

$$\begin{aligned} y_t &= c + u_t - \theta_1 u_{t-1} \\ &= c + (1 - \theta_1 L) u_t \end{aligned} \tag{1.27}$$

where c is a constant and $u_t \sim WN(0, \sigma^2)$. By computing the expected value of y_t

$$E[y_t] = c$$

can be seen that the constant term in (1.27) is the mean of the series, such that the MA(1) can be rewritten as

$$y_t = \mu + u_t - \theta_1 u_{t-1} \tag{1.28}$$

Since MA models are finite linear combinations of white noise sequences, they are always stationary. The same property can be deduced by calculating the moments of the series. We have already seen that the mean of a MA(1) is time invariant. By using the fact that u_t and u_{t-1} are uncorrelated, the variance of y_t is equal to

$$\text{Var}[y_t] = \sigma^2 + \theta_1^2 \sigma^2 = (1 + \theta_1^2) \sigma^2$$

which is also independent of time. Let us set $\mu = 0$ for simplicity and multiply each side of equation (1.28) by y_{t-h}

$$y_t y_{t-h} = u_t y_{t-h} - \theta_1 u_{t-1} y_{t-h}$$

by taking expectations, the autocovariances of the series are obtained

$$\gamma_h = \begin{cases} -\theta_1 \sigma^2 & \text{if } h = 1 \\ 0 & \text{if } h > 1 \end{cases}$$

Dividing γ_h by $\gamma_0 = Var[y_t]$, yields the ACF of y_t

$$\rho_0 = 1, \qquad \rho_1 = \frac{-\theta_1}{1+\theta_1^2}, \qquad \rho_h = 0 \text{ for } h > 1$$

As can be seen from this result, the ACF of a MA(1) is different from zero at lag 1 but is zero afterward, that is to say, it cuts off at lag 1. For a MA(2) model of the form

$$y_t = \mu + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} \tag{1.29}$$

the ACF satisfies

$$\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{1+\theta_1^2+\theta_2^2}, \qquad \rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}, \qquad \rho_h = 0 \text{ for } h > 2$$

such that it cuts off at lag 2. It is easy to see how this results extend to the general MA(q) model

$$\begin{aligned} y_t &= \mu + u_t - \theta_1 u_{t-1} - ... - \theta_q u_{t-q} \\ &= \mu + u_t - \sum_{i=1}^{q} \theta_i u_{t-i} \end{aligned} \tag{1.30}$$

which, setting $\theta(L) = 1 - \theta_1 L - ... - \theta_q L^q$, can equivalently be expressed as

$$y_t = \mu + \theta(L) u_t$$

The moments of a MA(q) are calculated as

- $E[y_t] = \mu$
- $\text{Var}[y_t] = \gamma_0 = (1 + \theta_1^2 + ... + \theta_q^2)\sigma^2$
- $\text{Cov}[y_t, y_{t-h}] = \gamma_h = \begin{cases} -(\theta_h + \theta_{h+1}\theta_1 + ... + \theta_q\theta_{q-h})\sigma^2 & \text{for } h = 1, ..., q \\ 0 & \text{for } h > q \end{cases}$

Hence, since the process (1.30) has constant mean, constant variance and a time invariant autocovariance structure, the conditions for weak stationarity are satisfied. Moreover, the autocovariances and consequently also the ACF of a MA(q) model can take on non-zero values only up to lag q, after which they have value zero, such that $\rho_h = 0$ for $h > q$. Since the ACF of a MA cuts off at the lag corresponding to its order, it can be used for identifying the order of the process. The parameters of a MA model are conventionally estimated via the maximum likelihood estimator. In order to build the likelihood function, two different approaches can be used: the conditional and the exact likelihood method. In the context of conditional likelihood, the initial disturbances, u_t for $t \leq 0$ are set equal to zero. The likelihood function is then calculated recursively using the expressions for the error terms at $t = 1, 2, ...$, yielding $u_1 = y_1 - \mu$, $u_2 = y_2 - \mu + \theta_1 u_1$ and so on. When using the exact likelihood method, instead, the initial value of the error terms is treated as an additional parameter and is estimated jointly with the MA coefficients. Both estimation procedures yield similar results for large sample sizes.

1.2.3 ARMA Processes

ARMA(1,1) Models

Autoregressive moving average (ARMA) models result from the combination of AR and MA models and are often used in financial applications for their capability of describing the dynamic structure of the data while relying on a limited number of parameters. An ARMA(1,1) has the form

$$y_t = \phi_0 + \phi_1 y_{t-1} + u_t - \theta_1 u_{t-1} \tag{1.31}$$

where $u_t \sim WN(0, \sigma^2)$, ϕ_0 is a constant term and $\phi_1 \neq \theta_1$. Assuming that y_t is a covariance stationary process, such that $E[y_t] = E[y_{t-1}] = \mu$, the mean of (1.31) is equal to

$$E[y_t] = \mu = \frac{\phi_0}{1 - \phi_1}$$

which is exactly the same as the mean of an AR(1) sequence from Section 1.2.1. Assuming for simplicity that $\phi_0 = 0$, the variance and autocovariances of the ARMA(1,1) can be calculated by means of the Yule-Walker equations

$$
\begin{aligned}
E[y_t y_t] &= \phi_1 E[y_{t-1}y_t] + E[u_t y_t] - \theta_1 E[u_{t-1}y_t] \rightarrow \ \gamma_0 = \phi_1\gamma_1 + \sigma^2(1+\theta_1^2 - \phi_1\theta_1) \\
E[y_t y_{t-1}] &= \phi_1 E[y_{t-1}y_{t-1}] + E[u_t y_{t-1}] - \theta_1 E[u_{t-1}y_{t-1}] \rightarrow \ \gamma_1 = \phi_1\gamma_0 - \theta_1\sigma^2 \\
E[y_t y_{t-2}] &= \phi_1 E[y_{t-1}y_{t-2}] + E[u_t y_{t-2}] - \theta_1 E[u_{t-1}y_{t-2}] \rightarrow \ \gamma_2 = \phi_1\gamma_1 \\
&\vdots \\
E[y_t y_{t-h}] &= \phi_1 E[y_{t-1}y_{t-h}] + E[u_t y_{t-h}] - \theta_1 E[u_{t-1}y_{t-h}] \rightarrow \ \gamma_h = \phi_1\gamma_{h-1}
\end{aligned}
$$

By substituting the expression for the lag-1 autocovariance γ_1 into the expression for γ_0, after some manipulation, we obtain the variance of the process y_t

$$\mathrm{Var}[y_t] = \gamma_0 = \frac{(1+\theta_1^2 - 2\phi_1\theta_1)\sigma^2}{1-\phi_1^2}$$

In order for the variance to be positive, it is required that $\phi_1^2 < 1$, which is equivalent to $|\phi_1| < 1$, the same condition that an AR(1) process must satisfy in order to be stationary. By looking at the expression for γ_1 it is clear that the first lag autocovariance of an ARMA(1,1) is not the same as that of an AR(1). However, the lag-2 autocovariance γ_2 is identical for both processes, and the same is also true for each following lag up to lag h. By calculating the autocorrelations of y_t, it can be seen that they satisfy

$$\rho_0 = 1, \qquad \rho_1 = \phi_1 - \frac{\theta_1\sigma^2}{\gamma_0}, \qquad \rho_h = \phi_1\rho_{h-1} \text{ for } h > 1$$

Hence, the ACF structure of an ARMA(1,1) is essentially equal to that of an AR(1) model, with the only difference that it starts its exponential decay at lag 2. Note that neither the ACF nor the PACF of an ARMA(1,1) process become zero at any finite lag.

ARMA(p, q) Models

An ARMA(p, q) model has the form

$$
\begin{aligned}
y_t &= \phi_0 + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 u_{t-1} - \ldots - \theta_q u_{t-q} + u_t \\
&= \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \theta_i u_{t-i} + u_t
\end{aligned}
\tag{1.32}
$$

By making use of the former notation for the AR and MA polynomials $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ and $\theta(L) = 1 - \theta_1 L - \ldots - \theta_q L^q$, (1.32) can be rewritten as

$$\phi(L) y_t = \phi_0 + \theta(L) u_t$$

From the fact the simple AR(p) and MA(q) models are special cases of the general ARMA(p, q) model, it follows that the latter enjoys some of the properties of both models. In particular, the stationarity of y_t entirely depends on the AR coefficients included in the ARMA model, such that (1.32) has the characteristic equation

$$1 - \phi_1 z - ... - \phi_p z^p$$

and is a weakly stationary process provided that all the roots of the characteristic equation are less than one in absolute value or equivalently, lie inside the complex unit circle. If at least one of the characteristic roots is equal to or greater than unity, y_t is said to be an autoregressive integrated moving average (ARIMA) model. The mean of a stationary ARMA(p, q) model is equal to

$$E[y_t] = \frac{\phi_0}{1 - \phi_1 - ... - \phi_p}$$

while the lag-h autocovariance and autocorrelation of the process satisfy

$$\gamma_h = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + ... + \phi_p \gamma_{h-p}$$
$$\rho_h = \phi_1 \rho_{h-1} + \phi_2 \rho_{h-2} + ... + \phi_p \rho_{h-p}$$

for $h > q$. This result does not hold as long as $h \leq q$ due to the correlation between the terms $\theta_h u_{t-h}$ and y_{t-h}. The ACF of an ARMA(p, q) starts its exponential decay at lag q, while the PACF begins going to zero from lag p. Both the ACF and the PACF do not cut off at any finite lag, such that they cannot be used for determining the order of the ARMA model. As for MA models, estimation of ARMA models is performed by means of maximum likelihood.

1.3 Nonstationary Time Series Models

Nonstationarity is a common feature of many economic and financial time series such as interest rates, foreign exchange rates and stock price series. Nonstationary time series are characterized by lacking the tendency of revolving around a fixed value or trend and by being long memory processes, implying that shocks, i.e. unforseen changes in the value of a variable, have a persistent effect on nonstationary data. Whereas the influence of innovations on a stationary process such as those introduced in Section 1.2 progressively fades away as time passes and eventually disappears, the effect of a shock on a nonstationary process does not necessarily decrease with time and is likely to endure till infinity. In fact, whereas the ACF of a stationary series goes to zero at an exponential rate (see Section 1.2), that of a nonstationary process decays at a far slower linear rate as the number of lags increases, implying that a shock will affect the process indefinitely, i.e. have a persistent effect. Moreover, nonstationary processes are likely to display serially autocorrelated and heteroskedastic errors. Several studies (Kim et. al (2002), Busetti and Taylor (2003), Cavaliere and Taylor (2006, 2007), among others) showed that the presence of autocorrelation and ARCH effects in the residual series of a model might cause the invalidity of standard asymptotic test which are derived under the assumption of $i.i.d$ Gaussian residual distribution. Another problem of using nonstationary data in statistical analysis is that the inference drawn from a nonstationary process might result invalid. If two time series contain a stochastic trend (Section 1.3.2), regressing them against each other might produce a so-called spurious regression (Section 4.1), which is charachterized by a high level of fit, as measured by the R^2 coefficient, and a low value of the Durbin-Watson statistic, indicating highly autocorrelated residual. Although the t-statistic from the regression output might indicate the existence of a significant relationship between the nonstationary variables, the resulting model is without any economical meaning as past innovations permanently affect the system. Furthermore, standard assumptions employed in regression analysis are invalid when applied to nonstationary data. When the error terms are not independent, estimates of the regression coefficients will be inefficient though unbiased, while forecasts built on the spurious regression may be inaccurate and significance tests on the regression coefficients misleading. In fact, traditional significance tests are likely to indicate that the null hypothesis of no relationship between two nonstationary variables should be rejected in favor of the acceptance of a spurious relation, even when the variables are generated by independent processes. Hence, it is possible that uncorrelated nonstationary variables that are not bound by any sort of causal relationship appear as if they were highly correlated in a regression analysis. Depending on whether a nonstationary process is difference- or trend-stationary, first differencing or de-trending the series will produce a stationary process, i.e. will remove the stochastic or deterministic trend contained in the series, as outlined in Section 1.3.2. Nonstationary time series are often termed integrated processes, according to the following definition

Definition 6 (Integrated Process). *$\{y_t\}$ is an integrated process of order 1, written $y_t \sim I(1)$, if its first difference is a stationary series. More in general, a nonstationary time series y_t is I(d) if differencing* *it d times results in a process that has no unit roots, such that* $\Delta^d y_t \sim I(0)$.

Since the root of the characteristic equation of a nonstationary I(1) series is unity, processes that are integrated of order 1 are also referred to as unit root processes. In the following section, three classes of I(1) models are introduced: the random walk, the random walk with drift and the trend-stationary process.

1.3.1 Random Walk

A random walk process has the form

$$y_t = y_{t-1} + u_t \tag{1.33}$$

where u_t is a stationary disturbance term, distributed as Gaussian white noise with mean zero and variance σ^2. Although (1.33) can be seen as a special case of an AR(1) model, it does not satisfy the condition for weak stationarity $|\phi_1| < 1$, since the coefficient of the term y_{t-1} is unity. By applying the lag operator, (1.33) can be rewritten as

$$y_t = Ly_t + u_t \rightarrow \ (1 - L)y_t = u_t$$

The characteristic equation of the process y_t is $1 - \lambda = 0$, which has root $\lambda = 1$. Because their characteristic equation has a unit root, random walk models belong to the class of unit root processes. By means of repeated substitution, y_t can be expressed as the sum between the initial observation y_0 of the series and the sequence of error terms

$$\begin{aligned} y_t &= u_t + u_{t-1} + ... + u_1 + y_0 \\ &= y_0 + \sum_{i=0}^{t-1} u_{t-i} \end{aligned} \tag{1.34}$$

Provided that y_0 is constant, the expected value and the variance of y_t are calculated as

$$E[y_t] = E[u_t] + E[u_{t-1}] + ... + E[u_1] + E[y_0] = y_0$$

$$\mathrm{Var}[y_t] = \mathrm{Var}[u_t] + \mathrm{Var}[u_{t-1}] + ... + \mathrm{Var}[u_1] = \sigma_u^2 t$$

Although $E[y_t]$ is time invariant, $\mathrm{Var}[y_t]$ is a linear function of time and the variance of the random walk process y_t increases as time passes, implying nonstationarity. In the instance that the autoregressive coefficient lies outside the complex unit circle, $|\phi_1| > 1$, the variance of the random walk process would grow exponentially with time. Considering the autocorrelation function of the series

$$\mathrm{Corr}[y_t, y_{t+h}] = \frac{t}{t\sqrt{1 + \frac{h}{t}}}$$

it is evident that it is also a function of the time index t. For any h, the correlation between y_t and y_{t+h} will go to one as $h \to \infty$, since h/t will decay to zero at a speed that depends on the value of t. Therefore, apart from being non-stationary, a random walk process is also not asymptotically uncorrelated. Moreover, a random walk process is not predictable nor mean-reverting, since for any forecast horizon, the point forecast of the random walk process is equal to the value of the process at the forecast origin t, as can be seen from the h-step ahead forecast of y_t

$$\hat{y}_t(h) = E[y_{t+h}|y_t, y_{t-1}, ...] = y_t, \forall h \geq 1$$

For this reason, the best prediction of a random walk model is always equal to the last observed value of the process. This phenomenon is due to the fact that, in each time period, the random walk y_t will wander up or down conditional on its previous value y_{t-1} with probability 0.5, which implies that, as long as the distribution of u_t is symmetrical around zero, the time path of the random walk process is entirely ruled by chance and is consequently not predictable. Otherwise stated, since the value of the random walk at present time t is obtained by adding an independent zero-mean random variable to its previous value y_{t-1}, the best forecast of y_t at any future point in time h is the value of y_t today. Consequently, a random walk process is likely to deviate strongly from its mean value and to cross it rarely, which is why their behavior is termed not mean-reverting. Like many nonstationary processes, the random walk is a long memory processes, implying that the effect of a shock on the series is persistent and does not die out with time, such that the process 'remembers' all past innovations.

1.3.2 Random Walk with Drift

The random walk with drift is formulated as

$$y_t = \mu + y_{t-1} + u_t \tag{1.35}$$

where $\mu = E[y_t - y_{t-1}]$ is the time trend of the series, which is also called the drift of the model. According to equation (1.34), the random walk with drift can be generalized to

$$\begin{aligned} y_t &= \mu t + u_t + u_{t-1} + ... + u_1 + y_0 \\ &= y_0 + \mu t + \sum_{i=0}^{t-1} u_{t-i} \end{aligned} \tag{1.36}$$

The sign of the slope constant μ rules the direction of the time path of the series, such that if $\mu > 0$, the value of the series increases until infinity as t increases, if $\mu < 0$ the series goes to $-\infty$, while the steepness of the movement is dictated by the magnitude of μ. By considering (1.36) and setting the starting value of the process y_t equal to zero, $y_0 = 0$, the expected value of y_t is

$$E[y_t] = \mu t$$

implying that the value of y_t will increase with time if the slope μ is positive, while it will decrease if the slope is negative. The h-step ahead forecast of y_t

$$\hat{y}_t(h) = E[y_{t+h}|y_t, y_{t-1}, ...] = \mu h + y_t$$

indicates that the best forecast of the process (1.35) for any forecast horizon h is the value of the process at the forecast origin t plus the drift term at time h, μh. The random walk and the random walk with drift are examples of difference-stationary processes, for they contain a stochastic trend whose changes are not fully predictable and which is represented by their cumulated errors $\sum_{i=0}^{t-1} u_{t-i}$, as it is shown in equations (1.34) and (1.36). In the case of the pure random walk (1.33), the stochastic trend can be removed by applying the first difference operator $\Delta = 1 - L$, yielding the stationary process $\Delta y_t = y_t - y_{t-1} = u_t$. For a random walk with drift such as (1.35), this procedure results in $\Delta y_t = \mu + u_t$, which is stationary around a constant mean.

1.3.3 Trend-Stationary Process

A trend-stationary process has form

$$y_t = \alpha + Dt + u_t \tag{1.37}$$

where u_t is distributed as Gaussian white noise with mean zero and variance σ^2. The process y_t is stationary around a linear time trend, which is captured by the deterministic term Dt. Contrary to stochastic trends, deterministic trends are predictable. Since u_t is a stationary process, (1.37) will display a trend-reverting behavior and the realizations of the process will fluctuate randomly around the trend without drifting too far away from it. The growth rate of the process (1.37) is governed by the value of the trend coefficient D. The mean of a trend-stationary process $E[y_t] = \alpha + Dt$ is not time invariant, whereas the variance $\mathrm{Var}[y_t] = \sigma^2$ is finite and does not depend on time. The deterministic trend in y_t can be removed by performing a linear regression, which would produce a covariance-stationary process.

Chapter 2

Box-Jenkins Model Selection

Box and Jenkins (1976) propose a series of guidelines for model selection which apply to AR, MA, ARMA and ARIMA models, described in Section 1.2. The popularity of the Box-Jenkins method stems from the evidence that simple models that rely on a limited number of variables outperform large and complex econometric models in several circumstances, e.g. in out-of-sample forecasts, as highlighted *inter alia* by Nelson (1972) and Cooper (1972). The underlying idea is that parsimonious parametrization, i.e. the construction of models that fit the data well while avoiding to incorporate unnecessary many coefficients, can prove advantageous in many applications. The Box-Jenkins methodology for economic modeling can be summarized by the following steps:

1. Stationarity: as a preliminary check, the stationarity of a time series can be assessed by visually inspecting its time plot. Otherwise, specific tests exist that determine whether a series is stationary or not, which are treated in Chapter 3. If non-stationarity is detected, it is possible to suitably transform the considered data-set so that it satisfies the assumption of covariance stationarity. Stationarity in the mean is usually achieved by either removing the deterministic trend or taking the first difference of the analyzed series with respect to time, depending on the nature of the non-stationarity (see Chapter 1.3). When a time series is highly volatile or when its variance is unstable, a logarithmic or power transformation might be appropriate (Pfaff (2006)).

2. Seasonality: if a time series is seen to present seasonal behavior, all seasonal patterns

should be removed from the data before a model is estimated. This is due to the fact that seasonality can be responsible for not otherwise justified high levels of volatility, which might distort the inference that is drawn from seasonally-unadjusted data. Seasonality can be removed by means of seasonal differencing (see Section 3.5) or by using appropriate seasonal models for describing the data-set.

3. Order specification: when it is not known in advance, the appropriate order of a model can be determined by examining the empirical autocorrelation and partial autocorrelation functions, as outlined in the next section. If the analyzed time series exhibits seasonality, the proposed procedures are also valid for determining the seasonal equivalent of the model order. Otherwise, information criteria such as the Akaike, Schwarz-Bayesian or Hannan-Quinn criterion can be used for order determination (Section 2.1).

4. Estimation: after a tentative model has been specified as described in step 3, the parameters of the models can be estimated by OLS or maximum likelihood (Section 2.2).

5. Diagnostic checking: this step consists in verifying whether the estimated model is adequate, in the sense that it describes the relevant features of the analyzed data-set sufficiently well. Many diagnostic tests are performed on the residual process of a time series model, since it is required that the residual series follows a white noise distribution in order for the selected model to be considered final. Hence, the most widespread tests are those aimed at detecting nonnormaity, ARCH effects or serial autocorrelation in the residual series. Specifically, residual normality is usually tested by means of the Jarque-Bera test (Section 2.3.1), whereas the null hypothesis of homoskedasticity is tested with the ARCH-LM test or equivalently with the Breusch-Pagan test (Section 2.3.2). In order to detect the presence of serial correlation in the residual process, the Box-Pierce and Ljung-Box tests (Section 2.3.3) are used. Alternatively, model checking can be performed via overfitting, a practice which consists of deliberately adding one or more extra coefficients to the model at some randomly selected lag. Overfitting ought not to affect the model greatly when it is adequate, such that the added coefficients should not appear to be statistically significant. In case that the tentative model proves inadequate, a different model is entertained and steps 3 through 5 repeated.

The following section describes how steps 3 through 5 of the Box-Jenkins methodology are implemented in the case of an AR model as well as for other types of models.

2.1 Order Specification for AR Models

2.1.1 Partial Autocorrelation Function

Consider the AR(1) model from (1.3): even though y_{t-2} does not appear in the expression for the AR(1), the terms y_t and y_{t-2} are correlated with autocorrelation coefficient ρ_2. The ACF is built so that it captures all such 'invisible' correlations between non-adjacent lags in the model specification. On the contrary, the partial autocorrelation between y_t and y_{t-h} has the feature that it eliminates the effect of any intermediate terms y_{t-1} through y_{t-h+1}, such that for an AR(1) model, the partial autocorrelation between y_t and y_{t-2} is zero. In order to see how the partial autocorrelation function (PACF) is formed, subtract the mean μ from every observation of the process $\{y_t\}$, so to obtain the new autoregressive series

$$y_t^* = \phi_{01} + \phi_{11} y_{t-1}^* + u_t \tag{2.1}$$

where $y_t^* = y_t - \mu$ and u_t is an error term which may or may not be a white noise process. Since in this case there are no intermediate values of y between y_t and y_{t-1}, ϕ_{11} represents both the autocorrelation and the partial autocorrelation coefficient. Let us extend (1.27) by including one more lag

$$y_t^* = \phi_{02} + \phi_{21} y_{t-1}^* + \phi_{22} y_{t-2}^* + u_t \tag{2.2}$$

In this case the coefficient ϕ_{22} measures the correlation between y_t and y_{t-2} after the effect of the intermediate term y_{t-1} has been controlled for and is therefore called the partial autocorrelation coefficient. The PACF is obtained by repeating this procedure for all additional h lags contained in the model. Since in an AR(p) model there is no direct correlation between y_t and y_{t-h} for $h > p$, the PACF for an AR(p) should cut off at lag p and be zero afterward, such that all partial correlation coefficients ϕ_{hh} are zero for $h > p$. It follows that since the PACF becomes zero after the lag corresponding to the order of the AR model is reached, it can be used to identify autoregressive processes by determining their order in case p is unknown.

2.1.2 Information Criteria

Alternatively, the order p of an AR model can be specified by means of information criteria. For a Gaussian AR(k) model, there are three likelihood-based information criteria available to this purpose: the Akaike (1974), the Bayesian-Schwarz (1978) and the Hannan-Quinn information criterion, respectively defined as

$$\mathrm{AIC}(k) = \ln(\tilde{\sigma}^2) + \frac{2k}{T} \tag{2.3}$$

$$\mathrm{BIC}(k) = \ln(\tilde{\sigma}^2) + \frac{k}{T}\ln(T) \tag{2.4}$$

$$\mathrm{HQIC}(k) = \ln(\tilde{\sigma}^2) + \frac{2k}{T}\ln(\ln(T)) \tag{2.5}$$

where $\tilde{\sigma}^2$ is the maximum likelihood estimate of the residual variance, $k = p + 1$ is the total number of estimated parameters and T is the sample size. The first term on the RHS of the equations is a measure of the goodness of fit of the AR(k), i.e. how well the estimated model fits the data, while the second term is the so-called penalty function of the corresponding criterion. This way, a model is penalized according to the number of parameters it contains, hence endorsing parsimonious models, which can describe the features of the data using a limited number of parameters. The reason why parsimonious models are preferred is because the residual sum of squares $\sum_{t=1}^{T} \hat{u}_t^2$ is inversely proportional to the number of degrees of freedom. An increase in the number of variables in the model provokes a reduction in the number of degrees of freedom, while the coefficient standard errors and consequently their confidence intervals will be larger. Although there is no evidence that one criterion is superior to the others in terms of absolute performance, the BIC and the HQIC enjoy better large sample properties than the AIC. For sample sizes that approach infinity, in fact, both the BIC and HQIC are asymptotically consistent, meaning that the order k suggested by these two criteria will converge almost surely in probability to the true order p. Instead, the AIC can sometimes be biased towards selecting a model that is overparametrized. As it can be seen from expressions above, the penalty term for each additional parameter included in the model is different for each information criterion, with the consequence that different values for the order of the AR model may be suggested depending on the information criterion used. Looking at (2.3) and (2.4), since $\ln(T)$ is bigger than 2, the BIC includes a bigger penalty term than does the AIC, so that it will tend to select a more parsimonious model than the AIC, while the HQIC will suggest an order that is in between. For AR model specification, the value of the chosen information criterion is computed for $k = 0, ..., p$, where p is a prespecified upper bound for the model order. The value of k is selected for which the value of the criterion is minimal. Intuitively, when coefficients without any explanatory power are added to the model, the value of the information criterion will increase.

2.2 Estimation of AR Models

Estimation of a specified AR(p) model is usually performed by means of the conditional least-squares method. Consider the AR(p) model from (1.18) and, conditioning on the first p values of the series, rewrite the process starting at the $(p+1)$-th realization, such that $t = p+1, ..., T$. The resulting model can be estimated via OLS, so that the fitted model is

$$\hat{y}_t = \hat{\phi}_0 + \hat{\phi}_1 y_{t-1} + ... + \hat{\phi}_p y_{t-p} \tag{2.6}$$

where $\hat{\phi}_i$ is the OLS estimate of ϕ_i and the residual process is

$$\hat{u}_t = y_t - \hat{y}_t$$

The variance σ^2 of the disturbance term is estimated as

$$\hat{\sigma}^2 = \frac{\sum_{t=p+1}^{T} \hat{u}_t^2}{T - 2p - 1} \tag{2.7}$$

Alternatively, the model (1.18) can be estimated via conditional likelihood. In that case, the estimates of the model coefficients $\hat{\phi}_i$ remain unaltered, while the residual variance (2.7) becomes

$$\tilde{\sigma}^2 = \frac{\hat{\sigma}^2(T - 2p - 1)}{T - p}$$

The goodness of fit of the fitted AR(p) model (2.6) can be measured by means of the R^2 coefficient, defined as

$$R^2 = 1 - \frac{\text{residual sum of squares}}{\text{total sum of squares}} = 1 - \frac{\sum_{t=p+1}^{T} \hat{u}_t^2}{\sum_{t=p+1}^{T} (y_t - \bar{y})^2}$$

where $\bar{y} = \sum_{t=p+1}^{T} y_t/(T-p)$ and $0 \leq R^2 \leq 1$. For a stationary series, a value of the R^2 statistic which is close to one indicates that the estimated model fits the data well. The shortcoming of the R^2 measure consists in the fact that it is an increasing function of the number of parameters, such that its value tends to increase as more coefficients are included in the model, independently of their explanatory power. This problem can be overcome by using a modified version of the R^2 coefficient which, although it is no longer between zero and one, takes into account the number of parameters employed in the estimated model. This measure is known as the adjusted R^2 or $\bar{R}^2$ and is calculated as

$$\bar{R}^2 = 1 - \frac{\hat{\sigma}^2}{\mathrm{Var}[\hat{y}_t]}$$

where $\mathrm{Var}[\hat{y}_t]$ is the sample variance of y_t and $\hat{\sigma}^2$ is the residual variance (2.7). Alternatively, the information criteria described in Section 2.1 can also be used as a measure of the goodness of fit.

2.3 Diagnostic Checking

Once the fitted model is obtained, it is appropriate to verify that it is adequate, which is accomplished by checking whether the residual series $\hat{u}_t$ follows a white noise distribution. In particular, the residuals should satisfy the i.i.d. assumption and should follow a normal distribution with mean zero and constant variance σ^2. A preliminary analysis for model adequacy can be accomplished by visually inspecting the correlogram of the residual process in order to detect outliers, inhomogeneous variance, structural breaks and, more in general, periods in which the model does not fit the data sufficiently well. Evidence that the estimated model is adequate for the data-set is found when the residual process is shown to be normally distributed, homoskedastic and not serially correlated. If, however, the error process does not obey the white noise assumption, the fitted model should be refined, e.g. by eliminating regressors with no explanatory power, or a different model should be considered for describing the considered data.

2.3.1 Jarque-Bera Normality Test

The Jarque-Bera (1987) test is aimed at detecting non-normality in the residual series of a model by comparing the third and the fourth moment, i.e. the skewness and kurtosis, of the analyzed distribution with those of a normal distribution. When the skewess and kurtosis of the residual distribution are consistent with the corresponding moments of a normal distribution, the null hypothesis of residual normality is not rejected. Hence, the test considers the null hypothesis of normality and verifies that $E[u_t^s]^3 = 0$ and $E[u_t^s]^4 = 3$ against $E[u_t^s]^3 \neq 0$ and $E[u_t^s]^4 \neq 3$, where

u_t^s are the standardized residuals of the true model such that $u_t^s = u_t/\sigma$. The Jarque-Bera test statistic

$$\text{JB} = \frac{T}{6}\left[T^{-1}\sum_{t=1}^{T}(u_t^s)^3\right]^2 + \frac{T}{24}\left[T^{-1}\sum_{t=1}^{T}(u_t^s)^4 - 3\right]^2$$

is asymptotically distributed as a χ^2 with two degrees of freedom if the null hypothesis is not rejected. A multivariate version of the Jarque-Bera test exists as well and it is used for testing the null hypothesis of normality on the residual series of VAR and VECM models, introduced in Section 4.3 and Section 4.4, respectively. Formally, the multivariate Jarque-Bera test is a generalization of its univariate counterpart, which is based on a standardization of the residual series by means of a Choleski decomposition of the residual covariance matrix $\boldsymbol{\Sigma}$, estimated as

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T}(\hat{\boldsymbol{u}}_t - \bar{\hat{\boldsymbol{u}}})(\hat{\boldsymbol{u}}_t - \bar{\hat{\boldsymbol{u}}})'$$

The standardized residuals are defined as

$$\hat{\boldsymbol{u}}_t^s = \frac{1}{\tilde{\boldsymbol{P}}}(\hat{\boldsymbol{u}}_t - \bar{\hat{\boldsymbol{u}}})$$

where $\tilde{\boldsymbol{P}}$ is a lower triangular matrix with positive diagonal such that $\tilde{\boldsymbol{P}}\tilde{\boldsymbol{P}}' = \tilde{\boldsymbol{\Sigma}}$ is the Choleski decomposition of the covariance matrix $\boldsymbol{\Sigma}$. The test statistics for the multivariate Jarque-Bera test is

$$\text{JB}_{\text{mv}} = s_3^2 + s_4^2$$

where $s_3^2 = T\boldsymbol{b}_1'\boldsymbol{b}_1/6$ and $s_4^2 = T(\boldsymbol{b}_2 - \boldsymbol{3}_k)'(\boldsymbol{b}_2 - \boldsymbol{3}_k)'/24$ are the multivariate skewness and kurtosis, which are asymptotically distributed as a χ^2 with K degrees of freedom under the null hypotesis of normality. The parameters $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are the third and fourth non-central moments of the distribution of the standardized residuals $\hat{\boldsymbol{u}}_t^s$, while $\boldsymbol{3}_k = (3, ..., 3)'$ is a vector with dimensions $(K \times 1)$. The test statistics JB_{mv} is asymptotically distributed as a $\chi^2(2K)$.

2.3.2 Heteroskedasticity Tests

The ARCH-LM test (Engle(1982)), is a Lagrange multiplier (LM) test used for detecting autoregressive conditional heteroskedasticity (ARCH) in the residual process of the estimated model. The ARCH-LM test is based on the auxiliary regression

$$\hat{u}_t^2 = \beta_0 + \beta_1\hat{u}_{t-1}^2 + ... + \beta_q\hat{u}_{t-q}^2 + e_t \tag{2.8}$$

where $\hat{u_t}$ is the OLS estimate of u_t. The null hypothesis that there are no ARCH effects up to lag q in the residual series, $H_0 : \beta_1 = ... = \beta_q = 0$, is tested against the alternative $H_1 : \beta_i \neq 0$ for $i = 1, ..., q$. The LM test statistics

$$\mathrm{ARCH}_{\mathrm{LM}} = TR^2$$

is computed using the coefficient of determination R^2 of the auxiliary regression (2.8) and the number of observations T. Under the null hypothesis, the LM test statistics follows an asymptotic χ^2 distribution with q degrees of freedom.

The presence of ARCH effects in the residual series of a VAR or VECM model (Sections 4.3 and 4.4) can be tested by means of the multivariate version of the ARCH-LM test for residual heteroskedasticity, which is based on the following auxiliary regression:

$$\mathrm{vech}(\hat{\boldsymbol{u}}_t \hat{\boldsymbol{u}}_t') = \boldsymbol{\beta}_0 + \boldsymbol{B}_1 \mathrm{vech}(\hat{\boldsymbol{u}}_{t-1} \hat{\boldsymbol{u}}_{t-1}') + ... + \boldsymbol{B}_q \mathrm{vech}(\hat{\boldsymbol{u}}_{t-q} \hat{\boldsymbol{u}}_{t-q}') + \varepsilon_t \tag{2.9}$$

where vech is the column-stacking operator for symmetrical matrices, which stacks the columns of a matrix from the main diagonal downward (Lüktepohl and Krätzig (2004)). The matrix $\boldsymbol{\beta}_0$ has dimensions $1/2(n(n+1))$ and the coefficient matrices $\boldsymbol{B}_i$ $(i = 1, ..., q)$ have dimensions $1/2(n(n+1)) \times 1/2(n(n+1))$, where n is the number of variables in a VAR(p) model. The null hypothesis of the multivariate ARCH-LM test is the same as that for the univariate test, namely that there are no ARCH effects in the residual process, such that the matrices $\boldsymbol{B}_i$ are jointly zero. Hence, the null hypothesis $H_0 : \boldsymbol{B}_1 = ... = \boldsymbol{B}_q = 0$ is tested against the alternative $H1 : \boldsymbol{B}_i \neq 0$ for $i = 1, ..., q$. The test statistic for the multivariate ARCH-LM test

$$\mathrm{VARCH}_{\mathrm{LM}}(q) = \frac{1}{2} Tn(n+1) R_m^2$$

with

$$R_m^2 = 1 - \frac{2}{n(n+1)} \mathrm{tr}(\hat{\boldsymbol{\Omega}} \hat{\boldsymbol{\Omega}}_0^{-1})$$

where $\hat{\boldsymbol{\Omega}}$ is the covariance matrix of the residuals from (2.9), follows a $\chi^2(qn^2(n+1)^2/4$ distribution.

2.3.3 Autocorrelation Tests

Additionally, the residual series should be examined for serial correlation, since serially correlated residuals usually signalize a systematic movement in the data that the model coefficients do not properly capture. The Portmanteau test (Box and Pierce (1970)) or the Ljung-Box statistic (Ljung and Box (1978)) are used to test the null hypothesis that residual autocorrelation at lags 1 to m is zero against the alternative that at least one of the autocorrelations is different from zero. Formally, the tests consider the null hypothesis $H_0 : \rho_1 = ... = \rho_m = 0$ and the alternative hypothesis $H_1 : \rho_i \neq 0$ for $i = 1, ..., m$. For the residuals of an AR(p), the Portmanteau statistic is

$$Q(m) = T \sum_{i=1}^{m} \hat{\rho}_i^2$$

where T is the sample size, m is the maximum number of lags and $\hat{\rho}_i = T^{-1} \sum_{t=i+1}^{T} \hat{u}_t \hat{u}_{t-i}$ are the sample autocorrelations of the residual process u_t. Since the prespecified maximum lag lenght m can affect the power of the test, simulation studies recommend selecting $m \approx ln(T)$ to improve the performance of the test (Tsay (2001)). Under the null hypothesis that all m autocorrelation coefficients are not significantly different from zero, $Q(m)$ follows asymptotically a χ^2 distribution with m degrees of freedom. For an AR(p) model, the number of degrees of freedom of the χ^2 is equal to $m - p$, where p is the number of AR coefficients included in the model. Ljung and Box proposed a modified version of the Portmanteau statistic, which has more power in finite samples:

$$Q^*(m) = T^2 \sum_{i=1}^{m} \frac{\hat{\rho}_i^2}{T - i}$$

Since as the sample size approaches infinity, the terms $(T + 2)$ and $(T - h)$ cancel out in the expression for $Q^*(m)$, the Box-Pierce and the Ljung-Box statistics are asymptotically equivalent. The decision rule for the $Q^*(m)$ test is to reject the null hypothesis if $Q^*(m) > \chi^2_\alpha$, with χ^2_α being the $100(1 - \alpha)$-th percentile from a χ^2 distribution with m degrees of freedom. Alternatively, when the p-value associated to the $Q^*(m)$ statistic is provided, H_0 is rejected when the p-value is less than or equal to the significance level α.

The absence of serial correlation in the residuals of a VAR or a VECM model can be tested by means of the multivariate Portmanteau test, which tests the null hypothesis $H_0 : E[\boldsymbol{u}_t \boldsymbol{u}'_{t-i}] = 0$ for $i = 1, ..., m$ against the alternative that at least one autocorrelation of the residual process is non-zero. The test statistic is defined as

$$Q_{mv}(m) = T \sum_{i=1}^{m} \mathrm{tr}(\hat{C}'_i \hat{C}_0^{-1} \hat{C}_i \hat{C}_0^{-1}) \tag{2.10}$$

where $\hat{C}_i = T^{-1} \sum_{t=i+1}^{T} \hat{\boldsymbol{u}}_t \hat{\boldsymbol{u}}_{t-1}$ is the sample autocorrelation of $\boldsymbol{u}_t$. An adjustment for the

$Q_{mv}(m)$ statistic exists, which has better small sample properties:

$$Q^*_{mv}(m) = T^2 \sum_{i=1}^{m} \frac{1}{T-j} \mathrm{tr}(\hat{C}_i' \hat{C}_0^{-1} \hat{C}_i \hat{C}_0^{-1}) \tag{2.11}$$

The test statistics (2.10) and the adjusted Portmanteau stastistics (2.11) are approximately distributed as a χ^2 with $(n^2(m-p))$ degrees of freedom, where p is the number of coefficients included in the VECM or VAR model and n is the number of variables. As in the univariate case, the power of the multivariate Portmanteau test is affected by the chosen lag length m.

Chapter 3

Unit Root and Stationarity Tests

3.1 Dickey-Fuller Test

Consider the following AR(1) model:

$$y_t = \phi y_{t-1} + u_t \tag{3.1}$$

for $t = 1, ..., T$ where $y_0 = 0$ and $u_t \sim \mathrm{GWN}(0, \sigma^2)$. In order to test whether y_t follows a pure random walk such that it is I(1), one can examine the root of the autoregressive polynomial $\phi(z) = 1 - \phi z = 0$ of (3.1) and see if it is equal to one. Formally, this can be accomplished by testing the null hypothesis $H_0 : |\phi| = 1$ against the alternative hypothesis $H_1 : |\phi| < 1$. This unit root test was developed by Dickey and Fuller (1979) and goes by the name of Dickey-Fuller (DF) test. Under the null hypothesis, the characteristic equation of the AR(1) model has a unit root and the process is non-stationary, while under the alternative hypothesis, the coefficient of the lagged term is less than one and stability is ensured. Assuming that $|\phi| = 1$ implies that it is appropriate to difference the series in order to induce stationarity. In the context of a unit root test, the conventional t-statistics for a regression model do not follow a t-distribution under the null hypothesis of non-stationarity. Since under H_0, $y_t \sim I(1)$, the central limit theorem does not apply and the t-statistic is not asymptotically distributed as a standard normal even for large sample sizes. Instead, it follows a non-standard distribution named the Dickey-Fuller distribution. In this context, the test statistic used is

$$t_{\hat{\phi}} = \frac{\hat{\phi} - 1}{s.e.(\hat{\phi})} \tag{3.2}$$

where $\hat{\phi}$ is the least squares estimator of ϕ

$$\hat{\phi} = \frac{\sum_{t=1}^{T} y_t y_{t-1}}{\sum_{t=1}^{T} y_{t-1}^2}$$

and

$$\text{s.e.}(\hat{\phi}) = \sqrt{\frac{\sum(\phi - \phi')^2}{n}}$$

is the estimated standard error of ϕ, where ϕ' is the predicted value of ϕ and *n* is the number of observations. For practical purposes, equation (3.1) can be rewritten as

$$\Delta y_t = \psi y_{t-1} + u_t \tag{3.3}$$

so that $\psi = \phi - 1$ and the test hypotheses become $H_0 : |\psi| = 0$ and $H_1 : |\psi| < 0$. In this case the test statistic is

$$t_{\hat{\psi}} = \frac{\hat{\psi}}{s.e.(\hat{\psi})} \tag{3.4}$$

The critical values for the tests statistics in (3.2) and (3.4) have been tabulated by Dickey and Fuller (1979) using Monte Carlo simulations. For large samples, critical values for different significance levels are displayed in the following table:

Significance level	1%	5%	10 %
Critical values	-3.43	-2.86	-2.57

The fact that the DF critical values are much bigger than standard normal critical values in absolute terms suggests that more evidence is needed in order to reject the null hypothesis of a unit root. If usual critical values were used in a DF test, for a 95% confidence interval H_0 would be rejected much more often than in 5% of the cases. The decision rule is to reject H_0 in favor of H_1 if $t_{\hat{\phi}} < c$, where *c* is the critical value for the preferred significance level. The DF test is also applicable to series that contain an intercept or a linear trend, for which the follwing test regressions are used:

$$\Delta y_t = c + \psi y_{t-1} + u_t \tag{3.5}$$

$$\Delta y_t = c + Dt + \psi y_{t-1} + u_t \tag{3.6}$$

The presence of a trend in the DGP should be controlled for prior testing, since trend stationary processes which are $I(0)$ around their time trend may be mistaken for unit root processes in a standard DF test. The test hypotheses are the same as those for the test regression (3.3), while different critical values and *t*-statistic are employed for (3.5) and (3.6)

Significance level	1%	5%	10 %
Critical value	-3.96	-3.41	-3.12

The critical values of the DF test for a trended series are larger than those employed for testing (3.3), so that more evidence against H_0 is needed for rejection if the regression equation includes a linear trend. This version of the DF test is only valid as long as the disturbance term of the test regression is distributed as Gaussian white noise, so that u_t is not autocorrelated. However, if the true data generating process is unknown, it is possible that u_t contains autoregressive or moving average components. In order to account for serial correlation in the error term u_t an alternative test, called the augmented Dickey-Fuller (ADF) test, is run on the regression equation

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \theta_i \Delta y_{t-i} + u_t \tag{3.7}$$

which is allowed to additionally include an intercept and a time trend if needed. The ADF test is built by including *p* lags of the first-differenced dependent variable into the standard DF test regression. This way, the assumption that u_t follows a stationary AR(*p*) process is supported, while any serial correlation that might have been present in (3.7) is eliminated. The null hypothesis of the ADF test is still $H_0 : |\psi| = 0$ and the same *t*-statistics and critical values as for the regular test can be employed. One important issue concerning the test regression in (3.7) is the choice of the optimal number of lags. By including too many lags in the model specification, the number of parameters to estimate increases with a consequent loss of degrees of freedom. Degrees of freedom get used up because the large number of estimated parameters provokes a reduction in the quantity of observations available. For each lag that is added to the regression, one observation is lost. This can affect the power of the test to reject the null hypothesis of a unit root, with the result that for a stationary process, H_0 will be rejected less frequently than it otherwise should. On the other hand, if too few lags are included, the autocorrelation in the error terms will not be fully removed, which will cause the results of the test to be biased. A good rule of thumb is to use the frequency of the data for selecting the number of lags, so that, if the data is monthly, 12 lags of y_t will be included in the model. Alternatively, one can use information criterion such as those introduced in Section 2.1 and include the number of lags for which the value of the given criterion is minimized. The number of lags can also be selected recursively by starting with a relatively large number of lags, say n, and checking if the t-statistics on the coefficient of the n-th lagged first-difference is significant for a given critical value. In case it isn't, the test regression is re-estimated including $n-1$ lags and the process is repeated until a lag such that its corresponding coefficient is significantly different from zero is found. In case that u_t contains MA terms, which can be seen as an AR process of infinite order, the inclusion of a larger number of lags is required. It is possible to verify if an optimal lag length has been reached by looking for evidence of autocorrelation or structural breaks in the plot of the regression residuals. Otherwise, the correlogram of u_t should appear to be white noise when all serial correlation has been eliminated. In case that the model under inquiry is suspected to contain more than one unit root, the DF test can be extended in order to control for the presence of multiple unit roots (Enders (1995)). This is accomplished by performing the DF test on ψ_1 from the following equation

$$\Delta^2 y_t = \psi_1 \Delta y_{t-1} + u_t \tag{3.8}$$

If $H_0 : \psi_1 = 0$ cannot be rejected, $y_t \sim I(2)$. However, if ψ_1 results to be significantly different from zero, the hypothesis that $y_t \sim I(1)$ should be tested on the following regression

$$\Delta^2 y_t = \psi_1 \Delta y_{t-1} + \psi_2 y_{t-1} + u_t \tag{3.9}$$

If it appears that $\psi_1 < 0$ but $\psi_2 = 0$ then a single unit root is present in (3.9), while if they are both non-zero, the series y_t is stationary. Although it is rare for economic time series to have an order of integration that exceeds one, equation (3.8) can be generalized so that it accounts for the presence of a number *k* of unit roots

$$\Delta^k y_t = \psi_1 \Delta^{k-1} y_{t-1} + u_t \tag{3.10}$$

Following the same logic as before, if the null hypothesis of *k* unit roots is rejected, a new test for $k-1$ roots should be performed on the next test regression and so on. The underlying pattern can be formulated as

$$\Delta^k y_t = \psi_1 \Delta^{k-1} y_{t-1} + \psi_2 \Delta^{k-2} y_{t-1} + ... + \psi_k y_{t-1} + u_t \tag{3.11}$$

The same procedure should be repeated until it is not possible to reject H_0 or if the series is shown to be stationary. Specific conditions such as the presence of structural breaks in the model specification as well as seasonally adjusted data series can bias the results of a unit root test against the rejection of H_0. Borderline stationary time series, that exhibit a value of ϕ close to one can be mistaken for unit root processes when they undergo a unit root test, especially for small sample sizes. This problem can be overcome by coupling a unit root test with a stationary test, such as the KPSS (Section 3.3), when a close to unit root series is suspected.

3.2 Phillips-Perron Test

The Phillips-Perron (PP) test, developed by Phillips and Perron (1988), relies on a non-parametric correction of the Dickey Fuller test statistics which accounts for weak dependence and heteroskedasticity of the error process, such that it is no longer required to be $i.i.d.$ Unlike the DF test, which assumes the disturbance to follow a white noise series, the error term is now relieved from the assumption of serial uncorrelation and homogeneous distribution. The following two regression equations are considered for testing the null hypothesis of a unit root on the coefficient ϕ

$$y_t = \widehat{\mu} + \widehat{\phi} y_{t-1} + \widehat{u}_t \tag{3.12}$$

$$y_t = \widetilde{\mu} + \widetilde{\lambda}(t - \frac{T}{2}) + \widetilde{\phi} y_{t-1} + \widetilde{u}_t \tag{3.13}$$

where $(\widehat{\mu},\ \widehat{\phi})$ and $(\widetilde{\mu}, \widetilde{\lambda}, \widetilde{\phi})$ are the OLS estimates of the regression coefficients. The test statistics, which are described in Phillips (1988) and Phillips and Perron (1988), are built so that the limiting distribution and the critical values of the ADF statistics apply to the PP test as well. By considering the long-run variance σ^2_{Tl} and the residual variance σ^2,

$$\sigma^2_{Tl} = \lim_{T\to\infty} \left[\frac{\sum_{t=1}^{T} E(u_t^2)}{T} \right]$$

$$\sigma^2 = \lim_{T\to\infty} \frac{E\left[\sum_{t=1}^{T} u_t\right]^2}{T}$$

can be seen that as long as u_t is an $\mathrm{i.i.d.}$ process, such that $\sigma^2 = \sigma^2_{Tl}$, the limiting distribution of the test statistic is independent from the parameters of the process generating u_t. The PP test statistic Z_t for testing the null hypothesis of $\phi = 1$ on equation (3.12) or (3.13) is

$$Z_t = \left(\frac{\hat{\sigma}^2_{Tl}}{\hat{s}^2}\right)^{\frac{1}{2}} t_{\hat{\phi}} - \frac{1}{2}\left(\frac{\hat{s}^2 - \hat{\sigma}^2_{Tl}}{\hat{s}^2}\right)\left(\frac{T \cdot s.e.(\hat{\phi})}{\hat{\sigma}^2_{Tl}}\right) \tag{3.14}$$

where $t_{\hat{\phi}}$ is the DF test statistic from (3.2), $\hat{s}^2$ is the sample variance of the residuals and the long-run variance $\hat{\sigma}^2_{Tl}$ is estimated as

$$\hat{\sigma}^2_{Tl} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2 + \frac{2}{T}\sum_{s=1}^{l} w_{sl} \sum_{t=s+1}^{T} \hat{u}_t \hat{u}_{t-s} \tag{3.15}$$

with $w_{sl} = 1 - s/(l+1)$. It is understood that whenever the error process satisfies the $i.i.d.$ assumption, the ADF *t*-statistics can be employed equivalently to the PP statistics, the latter being a transformation of the former. However, when considering autocorrelated residuals, the PP statistics will result valid under the conventional critical values from the ADF test. The advantage of the PP test procedure is that the inclusion of additional lags in the model specification is no longer required. Hence, the possibility of a loss of power resulting from the misspecification of the lag length does not subsist. Concerning the power of the PP test as compared to the ADF test, Phillips and Perron (1988) present simulation evidence that if the error term is generated by a MA(1), such that

$$u_t = \varepsilon_t + \theta\varepsilon_{t-1}$$

the power of the PP test is higher as long as $\theta > 0$. When the disturbance contains a negative moving average term, however, the PP test presents substantial size distortion for finite sample sizes. As a consequence, a true null hypothesis of $\phi = 1$ is rejected more frequently than the nominal size, i.e. 5% percent. The authors suggest using the ADF test when the presence of negative MA components is suspected, while using the PP test if the error term is believed to be $i.i.d.$ or to contain a positive MA.

3.3 KPSS Test

The KPSS test, developed by Kwiatkowski, Phillips, Schmidt and Shin (1992), belongs to the class of stationarity tests. In this context, the null hypothesis is that of stationarity, while under the alternative the series has a unit root. Performing both a stationarity and a unit root test in order to evaluate whether or not a series is integrated is especially useful in case of near unity roots, for which the DF test has low power. Moreover, the choice of stationarity as the null hypothesis may be beneficial as many unit root test only reject the null of integration when there is strong evidence in favor of the alternative. The test regression for the KPSS test

$$y_t = \lambda t + \mu_t + u_t \tag{3.16}$$

with

$$\mu_t = \mu_{t-1} + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

involves a deterministic trend, a random walk and a stationary disturbance, which may be heteroskedastic. The initial value μ_0 is constant and represents the intercept of the regression. The test is carried out by examining $H_0 : \sigma_\varepsilon^2 = 0$ against $H_1 : \sigma_\varepsilon^2 > 0$, so that (3.16) is trend stationary under the null. In case that $\lambda = 0$, y_t is level stationary since it fluctuates around the level μ_0. The *t*-statistic employed by the KPSS test is the Lagrange multiplier (LM)

$$LM = \frac{\sum_{t=1}^{T} S_t^2}{\hat{\sigma_u^2}} \tag{3.17}$$

where $S_t = \sum_{i=1}^{t} \hat{u_i}$, $t = 1, ..., T$. The residual $\hat{u_t}$ is obtained from regressing y_t on either both an intercept and a trend, or on a constant only, depending on whether the user is testing for trend or level stationarity. The term $\hat{\sigma_u^2}$ in the *t*-statistic is a consistent estimate of the long run variance of the error term, obtained by using the residuals $\hat{u_t}$ and a Bartlett window $w(s, l) = 1 - s/(l+1)$ as a weighting function

$$\hat{\sigma_u^2} = T^{-1} \sum_{t=1}^{T} \hat{u_t}^2 + 2T^{-1} \sum_{s=1}^{l} w(s, l) \sum_{t=s+1}^{T} \hat{u_t}\hat{u}_{t-1} \tag{3.18}$$

In case y_t is level stationary, such that $\lambda = 0$, it can be shown that the distribution of the LM statistics from (3.17) converges to a function of the Brownian motion called a standard Brownian bridge

$$LM \xrightarrow{d} \int_o^1 V_1(r) dr \tag{3.19}$$

where $V_1(r) = W(r) - rW(1)$ and $W(r)$ is a Wiener process for $r \in [0, 1]$. For testing the hypothesis of trend stationarity, the model employed is that of equation (3.16) and the residuals $\hat{u_t}$ are the result of regressing y_t on both a constant and a trend. The LM statistic can be shown to be asymptotically distributed as a second-level Brownian bridge

$$LM \xrightarrow{d} \int_o^1 V_2(r)dr \tag{3.20}$$

with $V_2(r) = W(r) + (2r - 3r^2)W(1) + 6r(r-1)\int_o^1 W(s)ds$. The critical values for both the level and trend stationarity version of the test are derived by simulation and are provided in Kwiatkowsky et al. (1992). The KPSS test is a one-sided right-tailed stationarity test, for which the decision rule at a significance level α is to reject H_0 if the LM statistics is greater than the $(1-\alpha)$ quantile from the asymptotic distribution (3.19) or (3.20).

3.4 Andrews-Zivot Structural Break Test

Formally, structural breaks are defined as unforeseen events that cause a shift in the structure of the data-set under analysis; an example of such events is the promulgation of a new legislation or the introduction of a new system of measurement for a given data series. A structural break can either have a permanent effect on the series, or it can affect the series only for a limited period of time. In the first case, the structural shift is usually modeled by a step dummy variable —that is, a variable which has value 0 before the break date and 1 afterwards, while in the second case it is modeled by a pulse dummy, which is 1 in correspondence with the break and 0 otherwise. The presence of a structural shift in a stationary series makes it harder to distinguish it from an integrated series (Pfaff (2006)). Perron (1989, 1990) found evidence of a loss in the power of DF type tests to reject a false null hypothesis when the analyzed series contained a structural break. Andrews and Zivot (1992) developed a unit root test which accounts for the presence of a structural break in the series under consideration. Under the null hypothesis, the time series $\{y_t\}$, $t = 1, ..., T$ is a random walk with drift

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

and has no structural break, while under the alternative hypothesis, $\{y_t\}$ is stationary around a linear trend and one endogenous break occurs at an unknown point in time in the trend function of the process. The Andrews-Zivot test considers the following three models developed by Perron (1989) for the null hypothesis:

$$\text{Model(A)}: \; y_t = \mu + dD(T_B)_t + y_{t-1} + u_t \tag{3.21}$$

$$\text{Model(B)}: \; y_t = \mu_1 + y_{t-1} + (\mu_2 - \mu_1)DU_t + u_t \tag{3.22}$$

$$\text{Model(C)}: \; y_t = \mu_1 + y_{t-1} + dD(T_B)_t + (\mu_2 - \mu_1)DU_t + u_t \tag{3.23}$$

where $D(T_B)_t$ is a pulse dummy variable that is equal to 1 if $t = T_B + 1$ and zero otherwise, DU_t is a step dummy which is 1 if $t > T_B$ and 0 otherwise and $1 < T_B < T$ is the point in time in which the structural break occurs. Model (A) accounts for a one time change in the levels of the series, Model (B) allows for a change in the growth rate, whereas Model (C) is a combination of both. The alternative hypothesis of trend-stationarity relies on the following models:

$$\text{Model(A)}: \; y_t = \mu_1 + \beta t + (\mu_2 - \mu_1)DU_t + u_t \tag{3.24}$$

$$\text{Model(B)}: \; y_t = \mu + \beta_1 t + (\beta_2 - \beta_1)DT_t^* + u_t \tag{3.25}$$

$$\text{Model(C)}: \; y_t = \mu_1 + \beta_1 t + (\mu_2 - \mu_1)DU_t + (\beta_2 - \beta_1)DT_t^* + u_t \tag{3.26}$$

where

$$DT_t^* = \begin{cases} t - T_B & \text{if } t > T_B \\ 0 & \text{otherwise} \end{cases}$$

Under the alternative hypothesis, Model (A) permits a one time break in the intercept of the trend function of magnitude $(\mu_2 - \mu_1)$, Model (B) allows for a break in the slope of the trend function of size $(\beta_2 - \beta_1)$, while Model (C) accounts for both a change in the levels and the growth rate of the process. The ADF-type test regressions for the Andrews-Zivot test are the following:

$$y_t = \hat{\mu}^A + \hat{\theta}^A DU_t(\hat{\lambda}) + \hat{\beta}^A t + \hat{\alpha}^A y_{t-1} + \sum_{i=1}^{k} \hat{c}_i^A \Delta y_{t-i} + \hat{u}_t \tag{3.27}$$

$$y_t = \hat{\mu}^B + \hat{\beta}^B t + \hat{\gamma}^B DT_t^*(\hat{\lambda}) + \hat{\alpha}^B y_{t-1} + \sum_{i=1}^{k} \hat{c}_i^B \Delta y_{t-i} + \hat{u}_t \tag{3.28}$$

$$y_t = \hat{\mu}^C + \hat{\theta}^C DU_t(\hat{\lambda}) + \hat{\beta}^C t + \hat{\gamma}^C DT_t^*(\hat{\lambda}) + \hat{\alpha}^C y_{t-1} + \sum_{i=1}^{k} \hat{c}_i^C \Delta y_{t-i} + \hat{u}_t \tag{3.29}$$

where

$$DU_t(\lambda) = \begin{cases} 1 & \text{if } t > T\lambda \\ 0 & \text{otherwise} \end{cases}$$

$$DT_t^*(\lambda) = \begin{cases} t - T\lambda & \text{if } t > T\lambda \\ 0 & \text{otherwise} \end{cases}$$

and $\lambda = T_B/T$ is the unknown break-point. The break-point is estimated as the point in time for which the null hypothesis of a unit root with drift is the most likely to be rejected. Otherwise stated, the value of λ is chosen which minimizes the *t* statistic from Perron (1989)

$$t_{\hat{\alpha}^i}(\lambda),\ i = A, B, C$$

since the decision rule is to reject H_0 when the value of the test statistic $t_{\hat{\alpha}^i}$ is smaller than the corresponding critical value. This minimizing value is defined as $\hat{\lambda}^i_{\text{inf}}$, such that the *t* statistic for the Andrews-Zivot test is

$$t_{\hat{\alpha}^i}[\hat{\lambda}^i_{\text{inf}}] = \inf_{\lambda \in \Delta} t_{\hat{\alpha}^i}(\lambda),\ i = A, B, C \tag{3.30}$$

where Δ is a closed subset of [0,1]. The critical values for the Andrews-Zivot test are displayed in Zivot and Andrews (1992).

3.5 HEGY Test for Seasonal Unit Roots

Until now we have considered processes for which integration is assumed to occur only at the zero frequency of their spectrum. When a time series displays a distinct seasonal pattern, unit roots at different frequencies may exist, i.e. roots other than 1 that are on the complex unit circle. The spectrum of a seasonal time series peaks at different seasonal frequencies $\omega = 2\pi j/S,\ j = 1, ..., S-1$, where *S* is the periodicity of the data. Consider the data generating process

$$y_t = y_{t-s} + u_t \tag{3.31}$$

with seasonal frequency *S*. When the roots of (3.31) have modulus equal to one, y_t is termed a seasonally integrated process which is characterized by stochastic seasonality. Seasonally integrated processes have similar properties as non-seasonal unit root processes in that, for instance, they have long memory causing shocks to have a permanent effect on the time series (Fuller (1976)). When stochastic seasonality is present, it is necessary to apply the seasonal difference operator

$$\Delta^S = (1 - L^S) = (1 - L)(1 + L + L^2 + ... + L^{S-1}) = \Delta S(L)$$

in order to induce stationarity. For quarterly data, the seasonal operator can be factored into

$$\begin{aligned}\Delta^4 &= (1-L^4) = (1-L)(1+L)(1+L^2) \\ &= (1-L)(1+L)(1-iL)(1+iL)\end{aligned} \tag{3.32}$$

which has four unit roots ± 1 and $\pm i$. The root 1 is termed the zero frequency root or non-seasonal unit root, the root -1 has semiannual frequency, while $\pm i$ are complex roots that have annual frequency. The roots $+i$ and $-i$ are considered as having both a yearly cycle since they are impossible to tell apart in quarterly data (Hylleberg et al. (1990)). The development of a testing procedure for seasonal unit roots has been attempted by Dickey, Hasza and Fuller (1984) and Ahtola and Tiao (1987). These tests have the drawback that integration is not accounted for at all seasonal frequencies and that the roots are required to have the same modulus under the alternative hypothesis. The HEGY test by Hylleberg et al. (1990) allows to test for the presence of individual unit roots at specific seasonal frequencies, under the assumption of stochastic or deterministic seasonality. The authors show that if the process $\{y_t\}$ is generated by a AR(p) such as (1.18) with $p \geq 4$, then the AR operator

$$\phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$$

can be factored into

$$\begin{aligned}\phi(L) = &(1-L)(1+L)(1-iL)(1+iL) \\ &- \pi_1 L(1+L)(1-iL)(1+iL) + \pi_2 L(1-L)(1-iL)(1+iL) \\ &+ (\pi_3 + \pi_4 L)(1-L)(1+L) - \phi^*(L)(1-L)(1+L)(1-iL)(1+iL)\end{aligned} \tag{3.33}$$

using (3.32), such that $\phi(L)$ has non-seasonal, semiannual or annual unit roots depending on whether $\pi_1 = 0$, $\pi_2 = 0$ or $\pi_3 = \pi_4 = 0$. The HEGY test employs the following test regression

$$\Delta_4 y_t = \mu_t + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{3,t-2} + \sum_{i=1}^{p-4} \phi_i^* \Delta_4 y_{t-i} + u_t \tag{3.34}$$

where the regressors are defined as

$$\begin{aligned}y_{1t} &= (1+L)(1-iL)(1+iL)y_t = (1+L+L^2+L^3)y_t \\ y_{2t} &= -(1-L)(1-iL)(1+iL)y_t = -(1+L+L^2-L^3)y_t \\ y_{3t} &= -(1-L)(1+L)y_t = -(1-L^2)y_t\end{aligned}$$

and the deterministic term $\mu_t = \pi_0 + \beta t + \sum_{j=1}^{3} \alpha_j D_{j,t}$ can be included in (3.34) in order to allow for the possibility of an intercept, a linear time trend or seasonal dummy variables in the test regression. In order to avoid perfect multicollinearity, only three seasonal dummies are considered for quarterly series. The lagged seasonal differences $\Delta_4 y_{t-i}$ ensure that all autocorrelation is removed from the error process, such that it is white noise. The null hypothesis of non-seasonal and seasonal unit roots is tested by estimating (3.34) by OLS and determining the significance of $\pi_i,\ i = 1, .., 4$. If it results that certain π's are zero, then the corresponding roots are on the complex unit circle. Hence, testing that $\pi_1 = 0$ and $\pi_2 = 0$ by means of a *t*-test is equivalent to testing for the presence of a unit root at the zero and semiannual frequency. The null hypothesis of annual roots $\pi_3 = \pi_4 = 0$ is tested jointly by means of an *F*-type test. If the analyzed series contains no seasonal unit roots, then $\pi_2 \neq 0$ and either $\pi_3 \neq 0$ or $\pi_4 \neq 0$, while stationarity entails that $\pi_i \neq 0$ for $i = 1, .., 4$. Hylleberg et al. (1990) and Franses and Hobijin (1997) tabulated the critical values for quarterly data, which depend on the specification of the test regression and the sample size.

Null Hypothesis	Polynomial	Root	Frequency	Cycles/Year
$\pi_1 = 0$	$(1 - L)$	+1	0 (non-seasonal)	0
$\pi_2 = 0$	$(1 + L)$	-1	π (bimonthly)	6
$\pi_3 = \pi_4 = 0$	$(1 + L^2)$	$\pm i$	$\pm\frac{\pi}{2}$ (four-monthly)	3
$\pi_5 = \pi_6 = 0$	$(1 + L + L^2)$	$-\frac{1}{2}(1 \pm \sqrt{3}i)$	$\pm\frac{2\pi}{3}$ (quarterly)	4
$\pi_7 = \pi_8 = 0$	$(1 - L + L^2)$	$\frac{1}{2}(1 \pm \sqrt{3}i)$	$\pm\frac{\pi}{3}$ (semiannual)	2
$\pi_9 = \pi_{10} = 0$	$(1 + \sqrt{3}L + L^2)$	$-\frac{1}{2}(\sqrt{3} \pm i)$	$\pm\frac{5\pi}{6}$	5
$\pi_{11} = \pi_{12} = 0$	$(1 - \sqrt{3}L + L^2)$	$\frac{1}{2}(\sqrt{3} \pm i)$	$\pm\frac{\pi}{6}$ (annual)	1

Table 3.1: HEGY test: summary table for monthly data

The HEGY test for seasonal unit roots has been generalized for time series with monthly frequency by Franses (1991) and Beaulieu and Miron (1993). If the considered data has periodicity

$S = 12$, the seasonal difference operator is decomposed as

$$\begin{aligned}\Delta_{12} = 1 - L^{12} = (1-L)(1+L)(1-iL)(1+iL)\times \\ \left[1+\frac{1}{2}(\sqrt{3}+i)L\right]\left[1+\frac{1}{2}(\sqrt{3}-i)L\right]\times \\ \left[1-\frac{1}{2}(\sqrt{3}+i)L\right]\left[1-\frac{1}{2}(\sqrt{3}-i)L\right]\times \\ \left[1+\frac{1}{2}(\sqrt{3}+i)L\right]\left[1-\frac{1}{2}(\sqrt{3}-i)L\right]\times \\ \left[1-\frac{1}{2}(\sqrt{3}+i)L\right]\left[1+\frac{1}{2}(\sqrt{3}-i)L\right]\end{aligned}$$

Assuming an AR(p) with $p \geq 12$ as the data generating process, the HEGY test regression is

$$\begin{aligned}\Delta_{12}y_t = \mu_t &+ \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{3,t-2} \\ &+ \pi_5 y_{4,t-1} + \pi_6 y_{4,t-2} + \pi_7 y_{5,t-1} + \pi_8 y_{5,t-2} + \pi_9 y_{6,t-1} \\ &+ \pi_{10} y_{6,t-2} + \pi_{11} y_{7,t-1} + \pi_{12} y_{7,t-2} + \sum_{i=1}^{p-12} \phi_i^* \Delta_{12} y_{t-i} + u_t\end{aligned} \tag{3.35}$$

where

$$\begin{aligned}y_{1t} &= (1+L)(1+L^2)(1+L^4+L^8)y_t \\ y_{2t} &= -(1-L)(1+L^2)(1+L^4+L^8)y_t \\ y_{3t} &= -(1-L^2)(1+L^4+L^8)y_t \\ y_{4t} &= -(1-L^4)(1-\sqrt{3}L+L^2)(1+L^4+L^8)y_t \\ y_{5t} &= -(1-L^4)(1+\sqrt{3}L+L^2)(1+L^4+L^8)y_t \\ y_{6t} &= -(1-L^4)(1-L^2+L^4)(1-L+L^2)y_t \\ y_{7t} &= -(1-L^4)(1-L^2+L^4)(1+L+L^2)y_t\end{aligned}$$

and $\mu_t = \pi_0 + \beta t + \sum_{j=1}^{11} \alpha_j D_{j,t}$ are the deterministic terms that can be included in (3.35). If $\pi_1 = 0$, the process $\{y_t\}$ has a unit root at the zero frequency, whereas if any $\pi_i,\ i = 2, ..., 12$ is equal to zero, the process is seasonally integrated. Critical values for monthly seasonal frequencies can be found in Franses and Hobijn (1997). Table 3.1 summarizes the null hypotheses and corresponding roots of the HEGY test with monthly data.

Chapter 4

Cointegration and Error Correction

4.1 Spurious Regression

The classic linear regression model, which is widely applied in economics, notoriously underlies the assumption of stationarity for the explanatory variables, establishing this way the consistency and asymptotic properties of the OLS estimator, among others. In particular, for a static regression model such as

$$y_t = \beta_0 + \beta_1 x_t + e_t \tag{4.1}$$

under the assumption of stationarity and weak dependence of $\{y_t, x_t\}$, proofs of consistency require less than perfect collinearity as well as contemporaneous exogeneity of the regressor with respect to the error process

$$E[u_t | x_t] = 0$$

In this case, the OLS estimator is said to be consistent

$$\text{plim}\, \hat{\beta}_1 = \beta_1$$

although it still might be biased. Equivalently, it can be stated that as time passes, the number of observations and hence the amount of information collected from the sample data increases, so that sample moments converge to their population values. Stationarity is hence an essential prerequisite for convergence of theoretical and empirical values, since this is only ensured when population moments are constant in time. This requirement is obviously not satisfied by time series that are trending. Conventional regression techniques have been widely used on time series that are non-stationary, since they sometimes can appear to be effective. Nonetheless, the problem of non-sense regressions (Yule (1926)) and spurious regressions (Granger and Newbold (1974)) may arise when making inference on a regression performed on non-stationary data. A spurious regression is established whenever a statistically significant linear correlation appears to exist between two integrated variables that are actually unrelated. As an example, consider the DGP

$$y_t = y_{t-1} + u_t \tag{4.2}$$

$$x_t = x_{t-1} + \varepsilon_t \tag{4.3}$$

where $u_t \sim \mathrm{IID}(0, \sigma_u^2)$ and $\varepsilon_t \sim \mathrm{IID}(0, \sigma_\varepsilon^2)$. Since y_t and x_t are uncorrelated random walks, the coefficient β_1 in the regression equation (4.1) would be expected to converge in probability to zero and so would the R^2 coefficient from the regression, because the series are independent. By testing the null hypothesis of $\beta_1 = 0$ against the alternative that $\beta_1 \neq 0$ at a 5% significance level, the *t*-statistic should appear significant only 5% of the times. Through simulation experiments, Granger and Newbold (1974) showed that the *t*-statistic on $\hat{\beta}_1$ instead yields a significant result much more often than the nominal significance level. Hence, when a spurious regression is analyzed, the regression of two unrelated series produces a significant *t*-statistic even though no relationship subsists. Spurious relationships are known to persist in large sample sizes, for which the null of independence is rejected more frequently than it is for small samples. The *t*-statistic in this context neither follows a *t* distribution nor is asymptotically distributed as a standard normal, as would be the case for cross-sectional data or in regressions with stationary time series. The *t*-statistic, in fact, would only converge in probability to a standard normal if the e_t's in equation (4.1) were a serially uncorrelated process with mean zero, while instead $\{e_t\}$ follows a random walk under $H_0 : \beta_1 = 0$, so that the test statistics goes to infinity as the sample size gets large (Wooldridge (2003)). When considering the limiting distribution of the *F*-statistics for $H_0 : \beta_0 = \beta_1 = 0$, it can be seen that it diverges from the conventional *F* distribution and that rejection rates increase with sample size. However, when testing for serial correlation in the residuals, the results usually show that the model is misspecified, as the autocorrelation test converges in probability to values that imply an autocorrelation of unity —that is to say, $\mathrm{Corr}[e_t, e_{t+1}] \to 1$ as t increases. Another hint that the regression under analysis might be spurious is provided by the Durbin-Watson (DW) statistic. When considering a regression that is not spurious, the DW statistic computed from the residuals of (4.1) typically converges in probability to a non-zero value, meaning that an actual relationship between the time series subsists (Phillips (1986)). Instead, if the regression is spurious, the DW statistic goes to zero as $T \to \infty$. A good rule of thumb for discriminating spurious from genuine regressions is to compare the value of the R^2 coefficient to that of the DW statistics. If it appears that $R^2 > \text{DW}$, the regression is likely to be spurious (Granger and Newbold (1974)). The same conclusions can be drawn when considering a multiple regression equation with $y_t \sim I(1)$, in which either all or only some of the independent variables are also I(1). The regression model will result to be spurious unless some specific equilibrium relationship requirements are met.

4.2 Cointegrated Economic Variables

Consider two processes $\{y_t\}$ and $\{x_t\}$ that are integrated of order *d*. In general, a linear combination $z_t = y_t - \beta x_t$ of these two series will yield a process that is also I(*d*) for any number β. However, it is possible that for some $\beta \neq 0$, $z_t = y_t - \beta x_t \sim I(d-b),\ b > 0$. In this case, y_t and x_t are said to be cointegrated and β is referred to as the cointegrating vector. Engle and Granger (1987) define cointegration in the following way:

Definition 7 (Cointegration)**.** *Consider some* $(n \times 1)$ *vector of variables* $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})'$*, whose components are all I(d). If there exists a nonzero vector* $\boldsymbol{\beta}$ *for which* $\boldsymbol{Z}_t = \boldsymbol{\beta}'\boldsymbol{Y}_t \sim I(d-b)$*,* $b > 0$*, then the components of the vector* $\boldsymbol{Y}_t$ *are said to be cointegrated of order (d, b), denoted* $\boldsymbol{Y}_t \sim CI(d, b)$

Typically, in economic literature only the case of $CI(1, 1)$ variables is treated, since most financial time series are individually integrated of order one. When considering two I(1) processes, $d = b = 1$ and the combination $\boldsymbol{Z}_t$ yields a stationary result when the variables are cointegrated. In the definition by Engle and Granger, however, any relationship producing a reduction of the order of integration of the single variables by some positive scalar *b* is considered to be a cointegrating relationship. Note that this definition also implies that variables with different orders of integration cannot be cointegrated. Suppose in fact that the first two elements of $\boldsymbol{Y}_t$, y_{1t} and y_{2t}, are respectively I(d_1) and I(d_2), with $d_2 > d_1$, then any linear combination of y_{1t} and y_{2t} will be I(d_2). It is important to note that the cointegrating vector $\boldsymbol{\beta}$ is not uniquely defined, since if the combination $\boldsymbol{\beta}'\boldsymbol{Y}_t$ is stationary, so is also $c\boldsymbol{\beta}'\boldsymbol{Y}_t$ for any scalar $c \neq 0$. In this case both $\boldsymbol{\beta}$ and $c\boldsymbol{\beta}$ are cointegrating vectors for $\boldsymbol{Y}_t$. Hence, if $\boldsymbol{Y}_t$ is an $(n \times 1)$ vector, there can be up to $h < n$ [1] linearly independent cointegrating vectors, where the number h is called the cointegrating rank of $\boldsymbol{Y}_t$. Moreover, any linear combination of the elements of $\boldsymbol{\beta}$ also constitutes a cointegrating vector itself. In order to preserve the uniqueness of $\boldsymbol{\beta}$, a normalization assumption must be imposed on the cointegrating vector by setting its first element β_1 equal to one, which yields $\boldsymbol{\beta} = (1, -\beta_2, ..., -\beta_n)'$ [2] . The cointegrating relationship is then equal to

$$\boldsymbol{\beta}'\boldsymbol{Y}_t = y_{1t} - \beta_2 y_{2t} - ... - \beta_n y_{nt} \sim I(0) \tag{4.4}$$

or equivalently

$$y_{1t} = \beta_2 y_{2t} + ... + \beta_n y_{nt} + u_t,\ u_t \sim I(0) \tag{4.5}$$

The intuition behind the concept of cointegration is that of the existence of a long-run equilibrium relationship to which integrated time series eventually converge. When taken individually, I(1) variables can drift arbitrarily far apart from each other since they have no tendency to frequently return to their initial or mean value and their unconditional variance increases as a function of time. When instead two or more variables are cointegrated, so that a stationary combination $\boldsymbol{Z}_t = \boldsymbol{\beta}'Y_t$ of them exists, they are bounded by a so-called cointegrating relationship which causes their association to be always restored in the long-run. Short-term deviations are however possible and they are captured by the term u_t in (4.5), called the disequilibrium error, which measures the distance of the system from equilibrium at any point in time. The disequilibrium error is expected to have value zero in the long term, as economic forces act in the

[1] The reason why the number of cointegrating vectors must be strictly less than the number of variables is discussed more in detail in Section 4.4

[2] The following example shows that, in the bivariate case, there can exist only a single cointegrating vector: let y_t and x_t be I(1) variables with two cointegrating parameters $\beta_1 \neq \beta_2$, such that the regression model

$$y_t = \beta_j x_t + u_{jt} \sim I(0),\ j = 1, 2$$

is formed. By subtracting the first from the second equation, we obtain

$$(\beta_2 - \beta_1)x_t = u_{1t} - u_{2t}$$

Since x_t is a I(1) vector, the left hand side of this equation is I(1), whereas the right hand side is I(0) since it is the difference of two stationary processes. This is a contradiction unless $\beta_1 = \beta_2$, so that $u_{1t} = u_{2t}$.

direction of eliminating short-run departures from equilibrium. Hence, for $u_t = 0$, the system (4.5) is in long-run equilibrium, denoted

$$y_{1t} = \beta_2 y_{2t} + ... + \beta_n y_{nt} \tag{4.6}$$

Furthermore, cointegration requires that the I(1) variables share a common stochastic trend, which cancels out in the stationary combination $\boldsymbol{Z}_t = \boldsymbol{\beta}'\boldsymbol{Y}_t$ producing the cointegrating relationship. To this regard, consider the I(1) vector $\boldsymbol{Y}_t = (y_{1t}, y_{2t})'$ generated from the random walk process

$$y_{jt} = y_{jt-1} + u_{jt} = \sum_{i=1}^{t} u_{ji} + y_{j0},\ j = 1, 2 \tag{4.7}$$

and suppose that $\boldsymbol{Y}_t$ is cointegrated with a normalized cointegrating vector $\boldsymbol{\beta} = (1,\ -\beta_2)'$. The stochastic trend in (4.7) is captured by the cumulated errors $\sum_{i=1}^{t} u_{ji}$. The cointegrating relationship $\boldsymbol{Z}_t = \boldsymbol{\beta}'\boldsymbol{Y}_t$ between the elements of $\boldsymbol{Y}_t$ can be expressed as

$$\begin{aligned} \boldsymbol{Z}_t = \boldsymbol{\beta}'\boldsymbol{Y}_t = \begin{pmatrix} 1 - \beta_2 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} &= y_{1t} - \beta_2 y_{2t} \\ = \sum_{i=1}^{t} u_{1i} + y_{10} - \beta_2 &\left(\sum_{i=1}^{t} u_{2i} + y_{20} \right) \end{aligned} \tag{4.8}$$

For the stochastic trend to cancel out in $\boldsymbol{Z}_t$, we need that $\sum_{i=1}^{t} u_{1i} = \beta_2 \sum_{i=1}^{t} u_{2i}$, such that the two stochastic terms cancel out, from which follows that

$$\boldsymbol{Z}_t = y_{10} - \beta_2 y_{20} \sim I(0)$$

In order to achieve cointegration, the time series vector $\boldsymbol{Y}_t$ must contain less stochastic trends than variables. Hence, if a $(n \times 1)$ vector $\boldsymbol{Y}_t$ is cointegrated with $h < n$ cointegrating vectors, there will be $n - h$ common unit root stochastic trends. In case the data generating process contains a deterministic trend, the cointegrating relationships cancels the stochastic but not the deterministic trends. As a result, $\boldsymbol{Z}_t = \boldsymbol{\beta}'\boldsymbol{Y}_t$ is trend stationary, i.e. stationary around a linear trend. By regressing the elements of the vector $\boldsymbol{Y}_t$ against each other

$$y_{1t} = \beta_2 y_{2t} + u_t \tag{4.9}$$

it can be shown that if y_{1t} and y_{2t} are cointegrated, such that u_t is a stationary mean zero variable, then the OLS estimator $\hat{\beta}_2$ of β_2 is superconsistent, in the sense that apart from being consistent it also converges to the true value at rate T. In this case the standard t-statistic for testing $H_0 : \beta_2 = 0$ versus $H_1 : \beta_1 \neq 0$ is is asymptotically normally distributed (see Stock (1987)). Although it may be tempting to first difference I(1) variables in order to induce stationarity, such a procedure is inadvisable as it would destroy any potentially existing information about the long-run relationship of the series. For this reason, economic modeling involving stable relationship dynamics is usually conducted using the raw time series data rather than the differenced series, which have no long term solution. In economic theory, there are several examples of models that presume the existence of an equilibrium relationship between variables. In the permanent income model, consumption is linked to income by a cointegrating relationship, while purchasing power parity implies cointegration between nominal exchange rates and foreign and domestic prices. The Fisher equation presupposes that nominal interest rates and inflation are cointegrated. Money demand and growth theory models, covered interest rate parity and the expectation hypothesis of the term structure, which requires nominal interest rates with different maturities to be cointegrated, are other examples of economic theories that rely on the assumption of a stable long term relationship between integrated time series.

4.3 Vector Autoregressive Models

Vector autoregressive (VAR) models are the multivariate analog of AR models for single time series and are commonly used to model the dynamic relationship between a number of time series variables. VAR models were made popular by Sims (1980) as an alternative to simultaneous equations structural models. In a VAR model, the current value of each variable composing the system depends on its own previous values and on past values of the other time series contained in the VAR, plus an error term. Since in an AR model the value of the dependent variable is determined only by its own lags and an error term, it can be seen as a restricted case of a VAR model. In the simplest case of a bivariate VAR(p) with no deterministic terms, the model is formed by two variables y_{1t} and y_{2t} and can be written in explicit form as

$$\begin{aligned} y_{1t} &= \beta_{10} + \beta_{11} y_{1t-1} + ... + \beta_{1p} y_{1t-p} + \alpha_{11} y_{2t-1} + ... + \alpha_{1p} y_{2t-p} + u_{1t} \\ y_{2t} &= \beta_{20} + \beta_{21} y_{2t-1} + ... + \beta_{2p} y_{2t-p} + \alpha_{21} y_{1t-1} + ... + \alpha_{2p} y_{1t-p} + u_{2t} \end{aligned} \tag{4.10}$$

where u_{1t} and u_{2t} zero-mean white noise disturbance terms with covariance matrix $\mathbf{\Sigma}_u$. The model (4.10) can be expanded to include a number n of variables, such that each element of

$\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})$ will have its own equation. Since writing such a model in extended form would be a tedious task, the more compact notation

$$\underset{n\times1}{\boldsymbol{Y}_t} = \underset{n\times1}{\boldsymbol{\beta}_0} + \underset{n\times n}{\boldsymbol{\beta}_1}\,\underset{n\times1}{\boldsymbol{Y}_{t-1}} + \underset{n\times n}{\boldsymbol{\beta}_2}\,\underset{n\times1}{\boldsymbol{Y}_{t-2}} + ... + \underset{n\times n}{\boldsymbol{\beta}_p}\,\underset{n\times1}{\boldsymbol{Y}_{t-p}} + \underset{n\times1}{\boldsymbol{u}_t} \tag{4.11}$$

can be used equivalently. To understand where this notation comes from, consider the case in which $p = 1$ so that each variable in the VAR only depends on the first lag of y_{1t} and y_{2t}

$$\begin{aligned} y_{1t} &= \beta_{10} + \beta_{11} y_{1t-1} + \alpha_{11} y_{2t-1} + u_{1t} \\ y_{2t} &= \beta_{20} + \beta_{21} y_{2t-1} + \alpha_{21} y_{1t-1} + u_{2t} \end{aligned} \tag{4.12}$$

The model (4.12) can be rewritten in matrix form as

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \alpha_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \tag{4.13}$$

or otherwise as

$$\underset{n\times1}{\boldsymbol{Y}_t} = \underset{n\times1}{\boldsymbol{\beta}_0} + \underset{n\times n}{\boldsymbol{\beta}_1}\,\underset{n\times1}{\boldsymbol{Y}_{t-1}} + \underset{n\times1}{\boldsymbol{u}_t} \tag{4.14}$$

where $n = 2$ since there are only two variables in the system. By setting $\boldsymbol{\beta}(L) = \boldsymbol{I}_n - \boldsymbol{\beta}_1 L - ... - \boldsymbol{\beta}_p L^p$, it is possible to write (4.11) in lag operator notation as

$$\boldsymbol{\beta}(L)\boldsymbol{Y}_t = \boldsymbol{\beta}_0 + \boldsymbol{u}_t \tag{4.15}$$

The stability of a VAR(p) model is assessed by considering the characteristic roots of the $(np\times np)$ coefficient matrix $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_p)$ from equation (4.11). Provided that

$$\det\left(\boldsymbol{I}_n - \boldsymbol{\beta}_1\lambda - ... - \boldsymbol{\beta}_p\lambda^p\right) \neq 0$$

has characteristic roots $|\lambda| \leq 1$, the VAR process is stable, meaning that it generates weakly stationary time series. One important feature of VAR models is that they typically assume that all variables are endogenous. Even though a VAR model can include exogenous terms, e.g. seasonal dummy variables, linear time trends or stochastic exogenous variables, it is common practice to regard all the variables in a VAR as being endogenous (Lütkepohl (2011)). Provided that all the variables in the system are identified, the VAR model can be estimated by employing OLS on each separate equation as long as no contemporary terms appear on the RHS of the model equation. In fact, estimation via OLS is feasible only provided that the RHS terms are known at present time t. VAR models are often employed for forecasting macroeconomic time series, as they arguably fare better than large-scale simultaneous equations models in out-of-sample forecasts (Sims (1980)). On the other hand, VAR models can generate a degrees of freedom problem as more lags and equations are added to the system. If the VAR contains p lags of a number n of variables, the amount of parameters to be estimated will be $n + pn^2$, plus the estimate of $\mathbf{\Sigma}$. When the number of estimated parameters is large, degrees of freedom will be quickly used up, especially if the sample size is small, causing large standard errors and consequently wide confidence intervals for the coefficients of the model. Another shortcoming of VAR models is that they are a-theoretical, i.e. their specification can be accomplished devoid of much information from economic theory about the relationship among the variables. Vector autoregressions are commonly employed in testing the hypothesis that one or more of the variables in a VAR equation do not *Granger cause* the others. The definition of Granger causality is given in Granger (1969), although other analogous definitions exist and can be found, among others, in Sims (1972). For a VAR of the type specified in (4.10) the null hypothesis may be tested that y_2 does not Granger cause y_1, denoted

$$E[y_{1t}|I_{t-1}] \neq E[y_{1t}|J_{t-1}] \tag{4.16}$$

where the term I_{t-1} represents all past information that is available about both y_1 and y_2, while J_{t-1} only contains information about y_1. The test is aimed at verifying whether past values of y_2 are useful in forecasting the current value of the process y_1, namely y_{1t}. Note that this methodology can be extended to a set of $n-1$ variables, $\boldsymbol{Y}_{2t} = (y_{2t}, ..., y_{nt})'$, such that the ability of each regressor of forecasting y_{1t} is tested singularly. As long as (4.16) is valid, past values of y_2 as well as past values of y_1 are of some relevance in predicting y_{t1}. From this argument it follows that the coefficients attached to the lags of y_2 in the equation for y_1 should be significantly different from zero, which can be tested by means of an *F*-test for joint significance. If y_2 does not Granger cause y_1, then all the $\alpha_{1j},\ j = 1, ..., p$ in (4.10) should be equal to zero. Granger causality is only applicable for testing the lagged causality between y_1 and y_2 and it does not extend to causal contemporary relationships between the two variables. It is hence to be interpreted only in terms of forecasting ability of the past values of a variable with respect to the current value of another. In a broader sense, Granger causality can be stated in terms of the distribution of y_1 and y_2, such that y_2 does not Granger cause y_1 if the distribution of

y_1, conditional on past values of both variables, is equal to the distribution of y_1, conditional on its past values only. In applications, instead of testing the entire distribution of y_1, it is common practice to examine only whether the conditional mean of y_1 depends on past values of y_2. Besides the already illustrated concept of Granger causality, the concept of instantaneous causality is often used in applied econometric works. Formally, there is instantaneous causality between two variables y_{1t} and y_{2t} if

$$E[y_{1t}|I_t \cup \{y_{2,t+1}\}] \neq E[y_{1t}|I_t]$$

where the term I_t represents the set of past and current information about y_{2t} and $y_{2,t+1}$ is the future value of y_{2t}. In other words, we say that y_2 instantaneously causes y_1 if a better forecast of y_1 is achieved, at time t, by adding future information about the value of y_2, $y_{2,t+1}$, to past and current information about y_2. The concept of instantaneous causality is a symmetric one, such that if y_2 instantaneously causes y_1, also y_1 instantaneously causes y_2 and the causal direction is unspecified (Lütkepohl (2006)). This is due to the fact that, if the error terms u_{2t} and u_{1t} of bivariate VAR model such as (4.10) are uncorrelated, then y_{2t} is instantaneously causal of y_{1t} and vice versa (Lütkepohl and Krätzig (2004)).

In order to choose the number of lags of y_1 and y_2 to be included in the model (4.10), an *F*-test can be performed separately on each equation forming the VAR. Alternatively, the lag length may be selected by means of the multivariate version of the information criteria by Akaike (1974), Schwarz (1978) and Hannan-Quinn, defined as

$$\mathrm{MAIC} = \log|\hat{\mathbf{\Sigma}}| + \frac{2k}{T} \tag{4.17}$$

$$\mathrm{MBIC} = \log|\hat{\mathbf{\Sigma}}| + \frac{k}{T}\log(T) \tag{4.18}$$

$$\mathrm{MQIC} = \log|\hat{\mathbf{\Sigma}}| + \frac{2k}{T}\log\left(\log(T)\right) \tag{4.19}$$

where $\hat{\mathbf{\Sigma}}$ is the estimated variance-covariance matrix of the residuals, $k = n + pn^2$ is the total number of regressors in a VAR with *n* equations and *p* lags of each variable and *T* is the number of observations. As with univariate information criteria, the number of lags to be included in the model is the one minimizing the value of the chosen information criterion. Note that if y_2 does not Granger cause y_1, any set of lagged y_2 that appears in the model will not be significantly different from zero.

4.3.1 Forecasting

Consider the VAR(p) model from equation (4.11) and assume that the parameters $\boldsymbol{\beta}_i$ $(i = 1, ..., p)$ are known. In this instance, the best linear predictor —that is, the predictor with minimum mean-squared errors (MSE), of $\boldsymbol{Y}_{T+h}$ is the h-step ahead forecast

$$\boldsymbol{Y}_{T+h|T} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{Y}_{T+h-1|T} + ... + \boldsymbol{\beta}_p \boldsymbol{Y}_{T+h-p|T} \tag{4.20}$$

with $\boldsymbol{Y}_{T+j|T} = \boldsymbol{Y}_{t+j}$ for $j \leq 0$ and T is the forecast origin. The corresponding h-step forecast error is

$$\begin{aligned} \boldsymbol{Y}_{T+h} - \boldsymbol{Y}_{T+h|T} &= \boldsymbol{u}_{T+h} + \boldsymbol{\Phi}_1 \boldsymbol{u}_{T+h-1} + ... + \boldsymbol{\Phi}_{h-1} \boldsymbol{u}_{T+1} \\ &= \sum_{k=0}^{h-1} \boldsymbol{\Phi}_k \boldsymbol{u}_{T+h-k} \end{aligned} \tag{4.21}$$

where the matrices $\boldsymbol{\Phi}_k$ are computed recursively following

$$\boldsymbol{\Phi}_k = \sum_{j=1}^{k} \boldsymbol{\Phi}_{k-j} \boldsymbol{\beta}_j \text{ for } j = 1, 2, .. \tag{4.22}$$

with $\boldsymbol{\Phi_0} = \boldsymbol{I}_n$ and $\boldsymbol{\beta}_j = 0$ for $j > p$. The MSE matrix of the the h-step forecast $\boldsymbol{Y}_{T+h|T}$ is

$$\boldsymbol{\Sigma}_y(h) = E[(\boldsymbol{Y}_{T+h} - \boldsymbol{Y}_{T+h|T})(\boldsymbol{Y}_{T+h} - \boldsymbol{Y}_{T+h|T})'] = \sum_{j=0}^{h-1} \boldsymbol{\Phi}_j \boldsymbol{\Sigma} \boldsymbol{\Phi}_j'$$

Since the forecast errors have expectation zero, the predictors are unbiased and therefore the MSE is the forecast error variance, which can be used for constructing confidence intervals (Lütkepohl (2006)). Under the assumption that $\boldsymbol{u}_t$ has a multivariate normal distribution such that $\boldsymbol{u}_t \sim \text{i.i.d.}N(0, \boldsymbol{\Sigma})$, the forecast errors are also normally distributed and a $(1-\alpha)100\%$ forecast interval for the elements of $\boldsymbol{Y}_{T+h|T}$ can be set up as

$$[y_{i,T+h|T} - c_{1-\alpha/2}\sigma_i(h),\ y_{i,T+h|T} + c_{1-\alpha/2}\sigma_i(h)]$$

where $y_{i,T+h|T}$ is the i-th element of $\boldsymbol{Y}_{T+h|T}$, $c_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of a standard normal distribution and $\sigma_i(h)$ is the square-root of the i-th diagonal element of $\boldsymbol{\Sigma}_y(h)$, i.e. the standard deviation of the h-step forecast error for the i-th element of $\boldsymbol{Y}_t$.

Now consider the instance in which the VAR(p) process from equation (4.11) is an estimated one, such that its parameters $\boldsymbol{\beta}_i$ are unknown. The MSE h-step predictor of $\hat{\boldsymbol{Y}}_{T+h}$ is

$$\hat{\boldsymbol{Y}}_{T+h|T} = \hat{\boldsymbol{\beta}}_0 + \hat{\boldsymbol{\beta}}_1 \hat{\boldsymbol{Y}}_{T+h-1|T} + ... + \hat{\boldsymbol{\beta}}_p \hat{\boldsymbol{Y}}_{T+h-p|T} \tag{4.23}$$

where $\hat{\boldsymbol{\beta}}_i$ $(i = 1, ..., p)$ are the estimated parameter matrices and $\hat{\boldsymbol{Y}}_{T+j|T} = \boldsymbol{Y}_{T+j}$ for $j \leq 0$. The forecast error is now defined as

$$\boldsymbol{Y}_{T+h} - \hat{\boldsymbol{Y}}_{T+h|T} = \sum_{k=0}^{h-1} \boldsymbol{\Phi}_k \boldsymbol{u}_{T+h-k} + \left(\boldsymbol{Y}_{T+h|T} - \hat{\boldsymbol{Y}}_{T+h|T}\right)$$

where $\left(\boldsymbol{Y}_{T+h|T} - \hat{\boldsymbol{Y}}_{T+h|T}\right)$ is the proportion of the error which is due to estimating the VAR(p) model. The covariance matrix of the forecast error is

$$\hat{\boldsymbol{\Sigma}}_y(h) = \boldsymbol{\Sigma}_y(h) + E[(\boldsymbol{Y}_{T+h} - \hat{\boldsymbol{Y}}_{T+h|T})(\boldsymbol{Y}_{T+h} - \hat{\boldsymbol{Y}}_{T+h|T})']$$

whereas the $(1-\alpha)100\%$ forecast interval for the components of $\hat{\boldsymbol{Y}}_{T+h|T}$ has the form

$$[\hat{y}_{i,T+h|T} - c_{1-\alpha/2}\hat{\sigma}_i(h),\ \hat{y}_{i,T+h|T} + c_{1-\alpha/2}\hat{\sigma}_i(h)]$$

4.3.2 Impulse Response Functions

Impulse response analysis is used to quantify the impact that an innovation on the impulse variable has on the response variable in a VAR(p) model. Recall from Section 4.3 the stability condition for a VAR(p) process such as (4.11). In Section 1.2.1, the Wold decomposition theorem for weakly stationary AR processes was introduced; likewise, a stable VAR(p) process can be represented in the form of an infinite moving average process

$$\boldsymbol{Y}_t = \boldsymbol{\Phi}_0 \boldsymbol{u}_t + \boldsymbol{\Phi}_1 \boldsymbol{u}_{t-1} + \boldsymbol{\Phi}_2 \boldsymbol{u}_{t-2} + ... \tag{4.24}$$

where $\boldsymbol{\Phi}_0 = \boldsymbol{I}_n$ and the matrices $\boldsymbol{\Phi}_k$ are computed by recursive substitution according to (4.22). The (i,j)-th element of the matrix $\boldsymbol{\Phi}_k$ is interpreted as the expected response of variable $y_{i,t+k}$ to a unit change of variable y_{jt}, k periods of time ago. In practice, the elements of $\boldsymbol{\Phi}_k$ represent the impulse responses of the elements of $\boldsymbol{Y}_t$ with respect to the shocks $\boldsymbol{u}_t$, following

$$\boldsymbol{\Phi}_{ij,k} = \frac{\delta y_{i,t+k}}{\delta u_{jt}} = \frac{\delta y_{it}}{\delta u_{j,t-k}} \text{ for } i,j = 1, ..., n$$

For a stable VAR(p), the accumulated effects of the impulses can be obtained by adding the matrices $\boldsymbol{\Phi}_k$ according to

$$\boldsymbol{\Phi} = \sum_{k=0}^{\infty} \boldsymbol{\Phi}_k = (\boldsymbol{I}_n - \boldsymbol{\beta}_1 - ... - \boldsymbol{\beta}_p)^{-1}$$

This way, the accumulated impact of a unit shock in variable j on variable i at time k is computed for $k = 1, 2, ...$ This interpretation of impulse responses is only valid as long as innovations occur in one variable at a time, such that shocks are independent. If, however, the elements of $\boldsymbol{u}_t$ are correlated, such that $\boldsymbol{u}_t$ has non-zero off-diagonal elements, it is possible that a shock in one variable is followed by another shock in a different variable. In that case, orthogonal innovations, which are not contemporaneously correlated, are preferred. Hence, in an orthogonal impulse response, $\boldsymbol{u}_t$ is a diagonal matrix and a change in one of the elements u_{it} of $\boldsymbol{u}_t$ has no effect on the other elements, since they are orthogonal. Orthogonal innovations are calculated by means of a Choleski decomposition of the covariance matrix of the error process $\boldsymbol{\Sigma}$. If $\boldsymbol{P}$ is a lower triangular matrix, the Choleski decomposition is such that $\boldsymbol{\Sigma} = \boldsymbol{P}\boldsymbol{P}'$ and the moving average representation of the VAR(p) (4.24) can be rewritten as

$$\boldsymbol{Y}_t = \boldsymbol{\Psi}_0\boldsymbol{\varepsilon}_t + \boldsymbol{\Psi}_1\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Psi}_2\boldsymbol{\varepsilon}_{t-2} + ...$$

with $\boldsymbol{\Psi}_0 = \boldsymbol{P}$, $\boldsymbol{\Psi}_i = \boldsymbol{\Phi}_i\boldsymbol{P}$ $(i = 1, 2, ...)$ and $\boldsymbol{\varepsilon}_t = \boldsymbol{P}^{-1}$. In orthogonal impulse responses, the elements of the matrix $\boldsymbol{\Psi}_i$ are interpreted as responses of the VAR(p) variables to the innovations $\boldsymbol{\varepsilon}_t$. Because the matrix $\boldsymbol{\Psi}_i$ is lower triangular, only a shock on the first variable in the system can have a simultaneous impact on all the other variables, whereas an innovation on the second variable will have no effect on the first variable, but only on the remaining $n - 2$ variables, and so on. The recursive ordering of the impulses can be represented as $y_{1t} \to y_{2t} \to ... \to y_{nt}$, such that changes in the variable on the left of the arrow affect all other variables in the system, but not the opposite. This fact implies that the outcome of an orthogonal impulse response analysis depends on the ordering of the variables. Therefore, a variable should be chosen as the first when it can reasonably be assumed that a change in that variable has an instantaneous impact on all the other variables in the VAR(p).

4.3.3 Forecast Error Variance Decomposition

The forecast error variance decomposition (FEVD) is used to determine the contribution of a shock in variable j to the h-step forecast error variance of variable i, based on the orthogonal impulse response matrices $\boldsymbol{\Psi}_k$ from Section 4.3.2. Recall the expression of the h-step forecast error for a VAR(p) with known coefficients from equation (4.21). Rewriting it in terms of the orthogonal innovations $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, ..., \varepsilon_{nt})'$ yields

$$\begin{aligned} \boldsymbol{Y}_{T+h} - \boldsymbol{Y}_{T+h|T} &= \boldsymbol{\Psi}_0 \boldsymbol{\varepsilon}_{T+h} + \boldsymbol{\Psi}_1 \boldsymbol{\varepsilon}_{T+h-1} + ... + \boldsymbol{\Psi}_{h-1} \boldsymbol{\varepsilon}_{T+1} \\ &= \sum_{k=0}^{h-1} \boldsymbol{\Psi}_k \boldsymbol{\varepsilon}_{T+h-k} \end{aligned} \tag{4.25}$$

where $\boldsymbol{\Psi}_i = \boldsymbol{\Phi_i P}$ $(i = 1, ..., p)$ and $\boldsymbol{P}$ is a lower triangular. The forecast error (4.25) of the i-th element of $\boldsymbol{Y}_{T+h}$, $y_{i,T+h}$, is expressed as

$$y_{i,T+h} - y_{i,T+h|T} = \sum_{k=0}^{h-1} \left(\psi_{i1,k}\ \varepsilon_{1,T+h-k} + ... + \psi_{in,k}\ \varepsilon_{n,T+h-k}\right)$$

Formally, the orthogonal innovations $\boldsymbol{\varepsilon}_t$ are serially uncorrelated and have variance one; therefore, the forecast error variance of the i-th element of $\boldsymbol{Y}_{T+h}$ has the form

$$\sigma_i^2(h) = \sum_{k=0}^{h-1} \left(\psi_{i1,k}^2 + ... + \psi_{in,k}^2\right) = \sum_{j=0}^{n} \left(\psi_{ij,0}^2 + ... + \psi_{ij,h-1}^2\right)$$

The term $\left(\psi_{ij,0}^2 + ... + \psi_{ij,h-1}^2\right)$ is interpreted as the contribution of shocks in variable j to the forecast error variance of the h-step forecast of variable i. Dividing this quantity by the forecast error variance $\sigma_i^2(h)$ yields the forecast error variance decomposition

$$\text{FEVD}_{ij}(h) = \frac{\left(\psi_{ij,0}^2 + ... + \psi_{ij,h-1}^2\right)}{\sigma_i^2(h)} \tag{4.26}$$

which is defined as the proportion, in percent, of the h-step forecast error variance of variable i which is due to variable j. Since the FEVD is based on orthogonal impulse responses, the ordering of the variables in vector $\boldsymbol{Y}_t$ influences the value of the FEVD.

4.4 Vector Error Correction Models

A cointegrated system is composed by time series whose time path is influenced by the magnitude of any deviation from long-run equilibrium. If disequilibrium is to occur in any time period, at least one of the cointegrated variables is required to change in response to this deviation in the following period, so that the system can return to long-term equilibrium. According to Granger representation theorem (Engle and Granger (1987)), any set of n, $n \geq 2$, I(1) time series that are cointegrated has an error correction representation and any set of time series that is error correcting is cointegrated. Hence, if an error correction representation exists, the variables must be cointegrated of order (1,1). A bivariate error correction (EC) model can be set up by augmenting a two variables VAR which is expressed in first differences as well as levels with the appropriate error correction terms. Moreover, specific restrictions must be placed on the coefficients of the VAR in order to enable the variables to be CI(1,1). To this regard, consider the VAR from (4.12), which we may rewrite without the intercept for simplicity

$$\begin{aligned} y_{1t} &= \beta_{11} y_{1t-1} + \alpha_{11} y_{2t-1} + u_{1t} \\ y_{2t} &= \beta_{21} y_{2t-1} + \alpha_{21} y_{1t-1} + u_{2t} \end{aligned} \tag{4.27}$$

where u_{1t} and u_{2t} are white noise disturbance terms. By making use of the lag operator, (4.27) can be formulated as

$$\begin{aligned} (1-\beta_{11}L) y_{1t} - \alpha_{11} L y_{2t} &= u_{1t} \\ -\alpha_{21} L y_{1t} + (1-\beta_{21}L) y_{2t} &= u_{2t} \end{aligned} \tag{4.28}$$

By solving (4.28) with respect to y_{1t} and y_{2t}, the solutions

$$y_{1t} = \frac{(1-\beta_{21}L)u_{1t} + \alpha_{11} L u_{2t}}{(1-\beta_{11}L)(1-\beta_{21}L) - \alpha_{11}\alpha_{21}L^2} \tag{4.29}$$

$$y_{2t} = \frac{\alpha_{21} L u_{1t} + (\beta_{11}L)u_{2t}}{(1-\beta_{11}L)(1-\beta_{21}L) - \alpha_{11}\alpha_{21}L^2} \tag{4.30}$$

are found. It can be seen that y_{1t} and y_{2t} have the same inverse characteristic equation $(1-\beta_{11}L)(1-\beta_{21}L) - \alpha_{11}\alpha_{21}L^2 = 0$, whose roots are obtained by solving the equation with respect to L. Letting $\lambda = 1/L$, after some manipulations the characteristic equation $\lambda^2 - (\beta_{11} + \beta_{21})\lambda + (\beta_{11}\beta_{21} - \alpha_{11}\alpha_{21}) = 0$ is obtained. Depending on the value of the characteristic roots (λ_1, λ_2), y_{1t} and y_{2t} may or may not be cointegrated of order (1,1). Consider for instance the case in which both λ_1 and λ_2 lie inside the unit circle. Under this circumstance, (4.29) and (4.30) yield stable results, so that both processes are stationary and cannot be CI(1,1). If, on the other hand, either one of the characteristic roots lies outside the unit circle, neither y_{1t} nor y_{2t} will be difference stationary, since they will both be integrated of an order bigger than one. In particular, if both characteristic roots are equal to unity, the two variables will be I(2). The necessary and sufficient condition for y_{1t} and y_{2t} to be CI(1,1), such that an error correction representation exists, is that one of the characteristic roots is unity, while the other one must be less than unity in absolute value. By setting $\lambda_1 = 1$, and multiplying (4.29) by $(1-L) = \Delta$, we obtain

$$\Delta y_{1t} = \frac{(1-\beta_{21}L)u_{1t} + \alpha_{11} L u_{2t}}{1-\lambda_2 L} \tag{4.31}$$

which is a stationary process for $|\lambda_2 < 1|$. The condition placed on the first characteristic root implies that

$$\lambda_1 = \frac{1}{2}\left(\beta_{11} + \beta_{21} + \sqrt{\beta_{11}^2 + \beta_{21}^2 - 4(\beta_{11}\beta_{21} - \alpha_{11}\alpha_{21})}\right) = 1$$

so that, after some simplifications, the coefficients of the VAR satisfy

$$\beta_{11} = \frac{1 - \beta_{21} - \alpha_{11}\alpha_{21}}{1 - \beta_{21}} \tag{4.32}$$

On the other hand, the constraint $|\lambda_2 < 1|$ entails that

$$\beta_{21} > -1 \tag{4.33}$$

and

$$\alpha_{11}\alpha_{21} + \beta_{21}^2 < 1 \tag{4.34}$$

since at least one of the coefficients α_{11} and α_{21} must be different from zero in order for y_{1t} and y_{2t} to cointegrate. By taking into account the coefficient restrictions (4.32), (4.33) and (4.34), the VAR model (4.27) can be rewritten as

$$\begin{aligned} \Delta y_{1t} &= (\beta_{11} - 1)y_{1t-1} + \alpha_{11}y_{2t-1} + u_{1t} \\ \Delta y_{2t} &= (\beta_{21} - 1)y_{2t-1} + \alpha_{21}y_{1t-1} + u_{2t} \end{aligned} \tag{4.35}$$

Considering that (4.32) entails that $\beta_{11} - 1 = -\alpha_{11}\alpha_{21}/(1 - \beta_{21})$, the system becomes

$$\begin{aligned} \Delta y_{1t} &= -[\alpha_{11}\alpha_{21}/(1 - \beta_{21})]y_{1t-1} + \alpha_{11}y_{2t-1} + u_{1t} \\ \Delta y_{2t} &= (\beta_{21} - 1)y_{2t-1} + \alpha_{21}y_{1t-1} + u_{2t} \end{aligned} \tag{4.36}$$

Provided that both α_{11} and α_{21} are non-zero, the cointegrating vector can be normalized[3] with respect to y_{1t}, yielding the error correction representation

$$\begin{aligned} \Delta y_{1t} &= \alpha_1(y_{1t-1} - \beta y_{2t-1}) + u_{1t} \\ \Delta y_{2t} &= \alpha_2(y_{1t-1} - \beta y_{2t-1}) + u_{2t} \end{aligned} \tag{4.37}$$

where $\alpha_1 = \frac{-\alpha_{11}\alpha_{21}}{1-\beta_{21}}$, $\beta = \frac{1-\beta_{21}}{\alpha_{21}}$ and $\alpha_2 = \alpha_{21}$. Now the model (4.37) is in error correction form and both variables y_{1t} and y_{2t} change in response to the disequilibrium of the previous period, captured by $y_{1t-1} - \beta y_{2t-1}$. Because Δy_{1t} and u_{1t} are stationary by assumption, it follows that the linear combination $y_{1t-1} - \beta y_{2t-1}$ must also be I(0), so that each equation in (4.37)

[3] Normalization entails that the coefficient of the variable that is being normalized is unity in the expression for the cointegrating vector.

is balanced in terms of order of integration. Hence, y_{1t} and y_{2t} are cointegrated with cointegrating vector $(1, -\beta)$. The coefficients α_1 and α_2 represent the speed of adjustment parameters, i.e. the proportion of disequilibrium that is eliminated within a time period. A large value of α_1 indicates that most of the previous period's disequilibrium has been corrected by y_{1t} in the following period. If, on the contrary, the value of α_1 is small, y_{1t} is not much responsive to deviations from long-run equilibrium. For the model in (4.37) to be error correcting, it is required that at least one of the speed of adjustment parameters is non-zero, which is ensured by the constraints (4.33) and (4.34). If one of the $\alpha_i, i = 1, 2$, happens to be zero, the corresponding variable does not help in restoring the long-run equilibrium relation and all the error correction is performed by the other variable. In fact, if we suppose that $\alpha_1 = 0$, $\Delta y_{1t} = u_{1t}$ and y_{1t} only changes in response to current shocks. Moreover, if $\alpha_1 < 0$ and $\alpha_2 > 0$, y_{1t} decreases and y_{2t} increases following a positive deviation from long-term equilibrium. Hence, if one of the speed of adjustment parameters has a negative sign, the value of the corresponding variable has been above equilibrium in the previous period and a downward adjustment is to be expected. If we suppose for instance that $y_{1t-1} > \beta y_{2t-1}$, then the value of y_{1t-1} has overshot equilibrium in $t-1$. In order for y_{1t} to return to its equilibrium value, it must be the case that $\Delta y_{1t} < 0$, which entails $\alpha_1 < 0$. The vector error correction (VEC) model in (4.37) can also be augmented by including an intercept and $p-1$ lagged changes of both the dependent and the explanatory variable into the model equation, resulting in

$$\begin{aligned}
\Delta y_{1t} &= \beta_{10} + \alpha_1(y_{1t-1} - \beta y_{2t-1}) + \sum_{i=1}^{p-1} \beta_{11}^i \Delta y_{1t-i} + \sum_{i=1}^{p-1} \alpha_{11}^i \Delta y_{2t-i} + u_{1t} \\
\Delta y_{2t} &= \beta_{20} + \alpha_2(y_{1t-1} - \beta y_{2t-1}) + \sum_{i=1}^{p-1} \beta_{21}^i \Delta y_{2t-i} + \sum_{i=1}^{p-1} \alpha_{21}^i \Delta y_{1t-i} + u_{2t}
\end{aligned} \tag{4.38}$$

With this notation, the additional requirements needed for the concept of Granger casuality (Section 4.3) to be applicable in the context of a VEC model can be easily illustrated. If the process y_{2t} does not Granger cause y_{1t}, then all the lagged terms Δy_{2t-i} should not appear in the equation for Δy_{1t}. This implies, in turn, that all the coefficients α_{11}^i attached to the lagged changes of y_{2t} are equal to zero, as would be the case for a bivariate VAR model expressed in first differences. When considering a VEC model such as (4.38), however, y_{1t} is additionally required to be irresponsive of any short-term disequilibrium when it is not Granger caused by y_{2t}, such that the speed of adjustment parameter α_1 is also equal to zero. In general, an *n*-

variable VEC model can be set up by letting $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})'$ be an $(n \times 1)$ vector of I(1) variables, such that the model has the form

$$\Delta \boldsymbol{Y}_t = \boldsymbol{\Gamma}_0 + \boldsymbol{\Pi} \boldsymbol{Y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{Y}_{t-1} + \boldsymbol{\Gamma}_2 \Delta \boldsymbol{Y}_{t-2} + ... + \boldsymbol{\Gamma}_p \Delta \boldsymbol{Y}_{t-p} + \boldsymbol{u}_t \tag{4.39}$$

where $\boldsymbol{\Gamma}_0$ is an $(n \times 1)$ vector of intercept terms, $\boldsymbol{\Pi}$ is an $(n \times n)$ matrix of parameters with at least one element $\pi_{jk} \neq 0$, $\boldsymbol{\Gamma}_i$ is an $(n \times n)$ coefficient matrix and $\boldsymbol{u}_t$ is an $(n \times 1)$ vector of disturbance terms. If the model (4.39) has an error correction form, there must be a linear combination of the I(1) components of $\boldsymbol{Y}_t$ that is stationary. By solving (4.39) with respect to $\boldsymbol{\Pi} \boldsymbol{Y}_{t-1}$, we obtain

$$\boldsymbol{\Pi} \boldsymbol{Y}_{t-1} = \Delta \boldsymbol{Y}_t - \boldsymbol{\Gamma}_0 - \sum_{i=1}^{p} \boldsymbol{\Gamma}_i \Delta \boldsymbol{Y}_{t-i} - \boldsymbol{u}_t \tag{4.40}$$

Since every element on the RHS of equation (4.40) is stationary, this must be the case for $\boldsymbol{\Pi} \boldsymbol{Y}_{t-1}$ also, such that each row of $\boldsymbol{\Pi}$ is a cointegrating vector of $\boldsymbol{Y}_t$. Notice that we imposed the constraint on the matrix $\boldsymbol{\Pi}$ that at least one of its elements is different from zero, such that its rank is also non-zero. This is because estimating the model (4.39) without the term $\boldsymbol{\Pi} \boldsymbol{Y}_{t-1}$ would yield an *n*-variable VAR model in first differences, with the consequence that $\boldsymbol{Y}_t$ would not adjust in response to the previous period's disequilibrium. It follows that if $\boldsymbol{\Pi}$ contains only zero elements, the model (4.39) would no longer be error correcting. At the opposite extreme, if we suppose $\boldsymbol{\Pi}$ to have full rank, i.e. $\text{rank}(\boldsymbol{\Pi}) = n$, then all the *n* variables contained in $\boldsymbol{Y}_t$ would be stationary and not cointegrated, since we consider only variables with an order of integration not bigger than 1. Hence, in order for an error correction representation of (4.39) to exist, the rank of $\boldsymbol{\Pi}$ must be equal to h, $0 < h < n$, such that there are h cointegrating vectors of the system. If the VAR process from (4.27) has unit roots, which is the case if the roots of $\det(I - \beta_1 \lambda) = 0$ lie outside the complex unit circle, than $\boldsymbol{\Pi}$ is a singular matrix, implying that it has reduced rank $h < n$. Consider to this instance the model (4.39) in restricted form

$$\underset{n \times 1}{\Delta \boldsymbol{Y}_t} = \underset{n \times n}{\boldsymbol{\Pi}} \underset{n \times 1}{\boldsymbol{Y}_{t-1}} + \underset{n \times 1}{\boldsymbol{u}_t} \tag{4.41}$$

The first element of the process $\boldsymbol{Y}_t$ has expression

$$\Delta y_{1t} = \pi_{11} y_{1t-1} + \pi_{12} y_{2t-1} + ... + \pi_{1n} y_{nt-1} + u_{1t} \tag{4.42}$$

We can normalize (4.42) with respect to y_{1t-1} so that we obtain

$$\Delta y_{1t} = \alpha_1 (y_{1t-1} + \beta_{12} y_{2t-1} + ... + \beta_{1n} y_{nt-1}) + u_{1t} \tag{4.43}$$

where $\alpha_1 = \pi_{11}$ and $\beta_{ij} = \pi_{ij}/\pi_{11}$. Equation (4.43) represents a vector error correction model with normalized cointegrating vector $\boldsymbol{\beta} = (1, \beta_{12}, ..., \beta_{1n})$ and speed of adjustment parameter α_1.

Chapter 5

Cointegration Tests

5.1 Engle and Granger Cointegration Test

Let $\boldsymbol{Y}_t = (y_{1t}, y_{2t})'$ be a vector of I(1) random variables and consider the bivariate static regression model

$$y_{1t} = \beta_2 y_{2t} + u_t \tag{5.1}$$

Engle and Granger (1987) propose a cointegration testing procedure based on performing a unit root test on the residual process from (5.1). In particular, the Engle and Granger (EG) test aims at verifying if $\boldsymbol{Y}_t$ is cointegrated with cointegrating vector $\boldsymbol{\beta} = (1, -\beta_2)'$, so that the linear combination $\boldsymbol{\beta'}\boldsymbol{Y}_t = y_{1t} - \beta_2 y_{2t} = u_t$ is I(0). The test can be implemented by applying a unit root test such as the ADF or PP test directly to the process u_t, provided that the cointegrating parameter $\boldsymbol{\beta}$ is prespecified and does not need to be estimated. In order to achieve cointegration, the error term should follow a stationary white noise process. Since the null hypothesis of the ADF test is that of unit root non-stationary data, the Engle and Granger (EG) test has the null of no cointegration. If the residual process is found to have a unit root, it is non-stationary and the random variables y_{1t} and y_{2t} do not cointegrate. In many applications, it can happen that the cointegrating vector $\boldsymbol{\beta}$ is unknown, in which case the EG procedure is performed using the OLS estimator $\hat{\beta}_2$ of β_2 from regression (5.1). A unit root test is then applied on the estimated residual

$$\hat{u}_t = y_{1t} - \hat{\beta}_2 y_{2t} \tag{5.2}$$

Since in this context an estimate of the actual cointegrating relationship is being used, the critical values for testing H_0 differ from those tabulated by Dickey and Fuller (1976). The correct critical values for a unit root test on the cointegrating residuals are obtained by Monte Carlo simulation and can be found in Engle and Granger (1987) and Engle and Yoo (1987). The critical values tabulated by Engle and Granger are suited for bivariate regression models with no more than 100 observations, whereas Engle and Yoo extend these results to fit systems containing up to 5 variables and for 50, 100 or 200 observations. Since under the null hypothesis of no cointegration we are estimating a spurious regression, the asymptotic distribution of the ADF test *t*-statistic is not anymore the usual DF distribution, but rather a function of Wiener processes. The correct distribution of the ADF and PP test is in fact the Phillips and Ouliaris distribution, named after its finders (see Phillips and Ouliaris (1990) for more details).

Despite the fact that, as previously mentioned, the OLS estimator $\hat{\beta}_2$ is superconsistent for β_2, Stock (1987) and Phillips (1991) found that when cointegration is given, $\hat{\beta}_2$ can be substantially biased in finite samples and is also not efficient. However, both problems can be solved by estimating β_2 using the dynamic OLS estimator $\hat{\beta}_{2,DOLS}$, as suggested by Stock and Watson (1993). The authors show that under certain conditions, the DOLS estimator $\hat{\beta}_{2,DOLS}$ is consistent, asymptotically normally distributed and efficient. To see how $\hat{\beta}_{2,DOLS}$ is built, consider a vector of random variables $\boldsymbol{Y}_t = (y_{1t}, \boldsymbol{Y}'_{2t})$ with $\boldsymbol{Y}_{2t} = (y_{2t}, ..., y_{nt})'$ and a normalized cointegrating vector $\boldsymbol{\beta} = (1, -\beta'_2)$. The cointegrating regression of y_{1t} on $\boldsymbol{Y}_{2t}$ is augmented with *p* leads and lags of $\Delta\boldsymbol{Y}_{2t}$ plus a deterministic trend D_t

$$y_{1t} = \gamma' D_t + \beta'_2 \boldsymbol{Y}_{2t} + \sum_{i=-p}^{p} \psi'_i \Delta\boldsymbol{Y}_{2t-i} + u_t \tag{5.3}$$

By estimating the augmented regression (5.3) by ordinary least squares, the DOLS estimator of β_2 is obtained. One important feature of the EG test is that, by construction, the null hypothesis of no cointegration is rejected whenever the dependent variable y_{1t} of the cointegrating regression cointegrates with at least one of the independent variables $y_{2t}, ..., y_{nt}$. Because the test is based on the assumption that there is one single cointegrating vector, it is unable to tell with how many regressors y_{1t} cointegrates. When considering only two variables, there can be at most one linear combination that is stationary, albeit if the system contains *n* variables, there can be up to *h* linearly independent cointegrating relationships, where $h < n$. The OLS regression approach employed in the EG test, however, is uncapable of finding more than one cointegrating relationship, independently of how many variables the estimated model contains. As we will see in the next section, a solution to this problem is proposed by Johansen (1988, 1992) and Stock and Watson (1988), who developed a methodology that allows for the appraisal of all h cointegrating vectors in an n-variate system. Since the EG cointegration test is based on the employment of unit root tests such as the ADF and the PP test, all shortcomings of the latter also affect the former.

5.2 Johansen Test

Johansen (1988) and Stock and Watson (1988) propose an alternative cointegration testing procedure which additionally allows to detect the presence of multiple cointegrating vectors. The problems involved with the estimation technique employed in the EG test do not arise in this context as both procedures provide for the use of the maximum likelihood estimator. Since the Johansen and the Stock and Watson methods are similar, we intend to treat only the former in this section. To see how the Johansen test is built, consider the VAR(p) model

$$\boldsymbol{Y}_t = \boldsymbol{\beta}_1 \boldsymbol{Y}_{t-1} + \boldsymbol{\beta}_2 \boldsymbol{Y}_{t-2} + \boldsymbol{\beta}_3 \boldsymbol{Y}_{t-3} + ... + \boldsymbol{\beta}_p \boldsymbol{Y}_{t-p} + \boldsymbol{u}_t \tag{5.4}$$

and subtract $\boldsymbol{Y}_{t-1}$ from each side of the equation

$$\Delta \boldsymbol{Y}_t = (\boldsymbol{\beta}_1 - \boldsymbol{I})\boldsymbol{Y}_{t-1} + \boldsymbol{\beta}_2 \boldsymbol{Y}_{t-2} + \boldsymbol{\beta}_3 \boldsymbol{Y}_{t-3} + ... + \boldsymbol{\beta}_p \boldsymbol{Y}_{t-p} + \boldsymbol{u}_t \tag{5.5}$$

then add $(\boldsymbol{\beta}_1 - \boldsymbol{I})\boldsymbol{Y}_{t-2}$ and subtract $(\boldsymbol{\beta}_1 - \boldsymbol{I})\boldsymbol{Y}_{t-2}$ from (5.5)

$$\Delta \boldsymbol{Y}_t = (\boldsymbol{\beta}_1 - \boldsymbol{I})\Delta \boldsymbol{Y}_{t-1} + (\boldsymbol{\beta}_2 + \boldsymbol{\beta}_1 - \boldsymbol{I})\boldsymbol{Y}_{t-2} + \boldsymbol{\beta}_3 \boldsymbol{Y}_{t-3} + ... + \boldsymbol{\beta}_p \boldsymbol{Y}_{t-p} + \boldsymbol{u}_t \tag{5.6}$$

again, add $(\boldsymbol{\beta}_2 + \boldsymbol{\beta}_1 - \boldsymbol{I})\boldsymbol{Y}_{t-3}$ and subtract $(\boldsymbol{\beta}_2 + \boldsymbol{\beta}_1 - \boldsymbol{I})\boldsymbol{Y}_{t-3}$ from the previous equation

$$\Delta \boldsymbol{Y}_t = (\boldsymbol{\beta}_1 - \boldsymbol{I})\Delta \boldsymbol{Y}_{t-1} + (\boldsymbol{\beta}_2 + \boldsymbol{\beta}_1 - \boldsymbol{I})\Delta \boldsymbol{Y}_{t-2} + (\boldsymbol{\beta}_3 + \boldsymbol{\beta}_2 + \boldsymbol{\beta}_1 - \boldsymbol{I})\boldsymbol{Y}_{t-3} + ... + \boldsymbol{\beta}_p \boldsymbol{Y}_{t-p} + \boldsymbol{u}_t \tag{5.7}$$

Continuing in the same manner recursively yields

$$\begin{aligned} \Delta \boldsymbol{Y}_t &= \boldsymbol{\Pi} \boldsymbol{Y}_{t-p} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{Y}_{t-1} + \boldsymbol{\Gamma}_2 \Delta \boldsymbol{Y}_{t-2} + ... + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{Y}_{t-p+1} + \boldsymbol{u}_t \\ &= \boldsymbol{\Pi} \boldsymbol{Y}_{t-p} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{Y}_{t-i} + \boldsymbol{u}_t \end{aligned} \tag{5.8}$$

where $\mathbf{\Pi} = \sum_{i=1}^{p} \boldsymbol{\beta}_i - \boldsymbol{I}$ and $\mathbf{\Gamma}_k = -\sum_{j=k+1}^{p} \boldsymbol{\beta}_j,\ k = 1, ..., p-1$, since the VAR($p$) parameters satisfy $\boldsymbol{\beta}_1 = \mathbf{\Gamma}_1 + \mathbf{\Pi} + \boldsymbol{I}$ and $\boldsymbol{\beta}_k = \mathbf{\Gamma}_k - \mathbf{\Gamma}_{k-1},\ k = 2, ..., p$. Note that the results of the Johansen test are sensitive to the number of lags that are included in the VEC model. In order to select the optimal lag length, the same procedures suggested for VAR models (Section 4.3) can be employed. Any evidence that the error terms in the estimated model are not distributed as white noise is a signal that the lag length is misspecified. The Johansen cointegration test focuses on the characteristic roots $\lambda_i,\ i = 1, ..., n$, of the matrix $\mathbf{\Pi}$ from the model (5.8), which are also called the eigenvalues of $\mathbf{\Pi}$ [1]. The rank of $\mathbf{\Pi}$ specifies the number of cointegrating vectors that are present in the system. Since the rank of a square matrix is equal to the number of its non-zero eigenvalues, the Johansen test is built so that it controls for the number of characteristic roots of $\mathbf{\Pi}$ that are non-zero in order to assess the number of independent cointegrating vectors in (5.8). If $\mathrm{rank}(\mathbf{\Pi}) = 0$, all of the elements of $\mathbf{\Pi}$ are zero and (5.8) becomes a VAR(p) model written in first differences instead of levels. In this instance, all of the $\Delta \boldsymbol{Y}_{it}$ sequences contain a unit root and none of the variables is cointegrated. If, on the contrary, $\mathrm{rank}(\mathbf{\Pi}) = n$, the vector process $\boldsymbol{Y}_t$ is stationary and also not cointegrated. In order for (5.8) to contain a single cointegrating vector, the rank of $\mathbf{\Pi}$ must be equal to 1, such that $\mathbf{\Pi}\boldsymbol{Y}_{t-p}$ represents the error correction term and (5.8) is a VEC model. In intermediate cases in which $1 < \mathrm{rank}(\mathbf{\Pi}) < n$, there are multiple cointegrating vectors. If we suppose that $\mathrm{rank}(\mathbf{\Pi}) = h$, such that the number of independent cointegrating vectors is h, then only these h linear combinations of the variables will yield a stationary result. The eigenvalues of $\mathbf{\Pi}$ can be ordered from the largest to the smallest, $\lambda_1 > \lambda_2 > ... > \lambda_n$, such that for $0 < \lambda_i < 1$ [2], λ_1 will be the closest to one and λ_n the closest to zero. Hence, if the rank of $\mathbf{\Pi}$ is zero, then $\lambda_i = 0,\ \forall i$ and the elements of $\boldsymbol{Y}_t$ are not cointegrated. The *t*-statistics employed in the Johansen test are the following:

$$\lambda_{\text{trace}}(h) = -T \sum_{i=h+1}^{n} \log(1 - \hat{\lambda}_i) \tag{5.9}$$

$$\lambda_{\max}(h, h+1) = -T \log(1 - \hat{\lambda}_{h+1}) \tag{5.10}$$

[1] A scalar λ is said to be an eigenvalue of a $(n \times n)$ square matrix $\mathbf{\Pi}$ if there exists a non-zero $(n \times 1)$ vector $\boldsymbol{c}$, such that $\mathbf{\Pi}\boldsymbol{c} = \lambda\boldsymbol{c}$. In this case the vector $\boldsymbol{c}$ is called the eigenvector that corresponds to the eigenvalue λ.

[2] This constraint ensures that $\mathbf{\Pi}\boldsymbol{Y}_{t-p}$ is stationary — that is to say, it has all roots inside the complex unit circle.

where $\hat{\lambda_i}$ is the estimated value of the *i*-th eigenvalue obtained from estimating the matrix $\mathbf{\Pi}$ and T is the number of usable observations. In the *t*-statistics (5.9) and (5.10), the expression $\log(1 - \lambda_i)$ is used to verify whether the eigenvalues of $\mathbf{\Pi}$ are significantly different from zero. Since $\log(1) = 0$, if $\lambda_i = 0$, then also $\log(1 - \lambda_i)$ will be equal to zero. If the i-th eigenvalue of $\mathbf{\Pi}$ happens to be non-zero, then $\log(1 - \lambda_i) < 0$, while $\log(1 - \lambda_j) = 0$ for all $j \neq i$. Similarly, if $\mathrm{rank}(\mathbf{\Pi}) = 1$, then $\log(1 - \lambda_1)$ will be negative while all λ_i will not be different from zero for $i > 1$. The λ_{trace} statistic tests the null hypothesis that the number of independent cointegrating vectors is less than or equal to h against the alternative that it is bigger than h. The test begins by considering *n* eigenvalues and recursively removes the largest until the number of eigenvalues which are significantly different from zero is found. If $\lambda_{\mathrm{trace}} = 0$, then all of the characteristic roots of $\mathbf{\Pi}$ are not significantly different from zero. The larger the value of λ_i is, the larger will be the λ_{trace} statistic. The $\lambda_{\max}$ statistic tests the null hypothesis that the number of cointegrating vectors is h against the alternative that it is $h + 1$. Similarly, if the value of λ_i is large, so will be the value of $\lambda_{\max}$. The $\lambda_{\max}$ test is conducted sequentially by testing the null hypothesis of an increasing number of cointegrating vectors. If the initial null hypothesis of $h = 0$ is rejected, the test proceeds by controlling for the presence of one cointegrating vector, $h = 1$, and so on until a value of h is found for which the null hypothesis can no longer be rejected. The critical values for both test statistics are obtained through simulation experiments and can be found in Johansen and Juselius (1990) or in Osterwald-Lenum (1992). According to these studies, the two statistics follow a non-standard distribution which depends on the number of nonstationary components under the null hypothesis and on whether a constant or a drift term are included in the model equations. However, for a single cointegrating vector, the asymptotic distribution of the Johansen test is the same as that of the EG test. In case that the model contains an intercept, the critical values for λ_{trace} and $\lambda_{\max}$ are larger, while they are smaller if the model is specified with a drift term. The null hypothesis of either test is rejected in favor of the alternative if the critical values are smaller than the corresponding *t*-statistic. The Johansen test additionally permits to test hypotheses on the coefficients of the cointegrating relationship by imposing restrictions on the matrix $\mathbf{\Pi}$, provided that these restrictions are not binding. In fact, as long as restricting $\mathbf{\Pi}$ does not affect the model too much, the number of cointegrating vectors should remain unaltered. To this purpose, the matrix $\mathbf{\Pi}$ is defined as the product of two $(n \times h)$ matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}'$, where h is the rank of $\mathbf{\Pi}$, which satisfy $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$. The matrix

$\boldsymbol{\beta}' = (1, \beta_2, ..., \beta_n)$ contains the cointegrating vectors, while $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)$ is the matrix of the speed of adjustment parameters and indicates the weight of each cointegrating vector entering the VAR(p) equations. Equation (5.8) can be rewritten using the new notation for $\boldsymbol{\Pi}$ as

$$\Delta \boldsymbol{Y}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{Y}_{t-p} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{Y}_{t-i} + \boldsymbol{u}_t \tag{5.11}$$

Via maximum likelihood, the VAR model (5.4) can be estimated as a VEC model, the rank of $\boldsymbol{\Pi}$ can be assessed and a value of $\boldsymbol{\alpha}$ can be chosen such that $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ holds. When estimating the model, it is possible to include a constant or a linear trend in the cointegrating vector, and\or add a drift term to the VEC model. If, for example, we include a constant c in (5.11), the VEC model becomes

$$\Delta \boldsymbol{Y}_t = \boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{Y}_{t-p} + c) + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{Y}_{t-i} + \boldsymbol{u}_t$$

whereas adding to the model both a drift vector μ_0 and a linear trend term Dt yields

$$\Delta \boldsymbol{Y}_t = \mu_0 + \boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{Y}_{t-p} + Dt) + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{Y}_{t-i} + \boldsymbol{u}_t$$

After the matrices $\boldsymbol{\alpha}$ and β' have been specified, restrictions can be tested on them. In this instance, the test statistic to be used is

$$\lambda_{\text{res}} = T \sum_{i=1}^{h} [\log(1 - \lambda_i^*) - \log(1 - \lambda_i)] \tag{5.12}$$

where λ_i^* and $\lambda_i,\ i = 1, ..., n$, are respectively the characteristic roots of the restricted and unrestricted model. The test statistic (5.12) is asymptotically distributed as a χ^2 with $h(n - m)$ degrees of freedom, where m is the number of restrictions placed on the matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}'$. Suppose we wish to test restrictions on $\boldsymbol{\beta}'$: if the value of (5.12) is bigger than the corresponding value from a χ^2 table, with degrees of freedom equal to the number of restrictions imposed on $\boldsymbol{\beta}$, then the restrictions enclosed in H_0 are binding and the null hypothesis should be rejected. When this is the case, the value of λ_i should appear to be large relative to λ_i^*, which indicates that the number of cointegrating vectors has changed, and in particular, that it has diminished in the restricted model. In the exact same manner it is possible to test restrictions on $\boldsymbol{\alpha}$.

Chapter 6

Empirical Results

This section is intended to provide an empirical implementation of the theories and methods introduced in the previous chapters. The focus of this survey is set on assessing whether a long-run relationship between the U.S. stock market index and a set of selected macroeconomic variables exist. In particular, the methodology of cointegration analysis is used in an attempt to capture the long-term linkages between the Dow Jones Industrial Average (DJX) and the set of macroeconomic variables composed by industrial production (IP), narrow money supply (M1), short-term interest rates (TB3M), crude oil price (OIL) and the consumer price index (CPI) as a proxy for the rate of inflation. To this purpose, the order of integration of the variables is first of all determined by means of an ADF, PP and KPSS unit root and stationarity test and it is ensured that all the series are free from structural breaks and seasonal integration (Section 6.2). The existence of a cointegrating relationship between the financial time series is established using Engle and Granger's cointegration test and the cointegrating rank of the system of variables is determined via Johansen's likelihood ratio tests (Section 6.3). The dynamic dependencies among the variables are modeled by a cointegrated multivariate VECM, whose adequacy for describing the DGP of the analyzed data-set is assessed by performing a series of standard asymptotic tests on the residual matrix of the model (Section 6.4). Finally, the short-term linkages between the selected macroeconomic variables and the stock index are examined by performing an impulse response analysis and the adequacy of the estimated VECM as a forecasting tool is inquired in the context of an out-of-sample forecast (Section 6.5).

6.1 Description of the Data

The data-set employed in this survey consists of monthly observations of five economic variables — that is, industrial production, money supply, interest rates, crude oil price and inflation rate, and the Dow Jones stock market index, collected during the period from January 1989 to August 2015, for a total of 320 observations for each time series. The choice of the observation frequency is dictated by the fact that daily or weekly data tend to fluctuate more and contain more noise than monthly data. According to Hakkio and Rush (1991), moreover, the consistency of results from cointegration analysis is guaranteed by choosing a long enough observation period rather than by increasing the frequency of the data. The macroeconomic variables are obtained from the Organization for Economic Co-operation and Development (OECD) database, except for crude oil price which was obtained from the database of the Economic Research Department of the Federal Reserve Bank of St. Louis[1]. The time series of the Dow Jones Industrial Average is obtained from the database Bloomberg. Industrial production, which proxies for real economic activity, is expressed as a volume index, seasonally adjusted, with OECD reference year 2012. Money supply is defined as the narrow money supply M1, seasonally adjusted, measured in billions of dollars. The secondary market rate for 3-month Treasury bills is used as a for proxy for short-term interest rates and is expressed in percent, not seasonally adjusted. The price for crude oil is the West Texas Intermediate, expressed in dollars per barrel and not seasonally adjusted, while the inflation rate is measured as the Consumer Price Index for all items, reference year 1982–1984, seasonally adjusted. Coherently with the post-Keynesian assumption of endogenous money supply advocated by Kaldor (1982, 1985), Palley (1982, 1994) and Moore (1979, 1983, 1986, 1988, 1989), this study treats narrow money supply as an endogenous variable[2]. Based on previous studies and economic theory, the following paragraph illustrates what impact the selected macroeconomic variables are expected to have on the stock market.

[1]The macroeconomic variables can be downloaded from http://stats.oecd.org/ and https://research.stlouisfed.org/.

[2]The post-Keynesian school of monetary thought maintains that money supply is endogenously driven by the decisions of commercial banks on liquidity reserves in response to the economy-wide demand for bank credit. Evidence in favor of the endogenous determination of money supply was found in different countries by a number of empirical studies (Kaldor (1982), Moore (1983), Pollin (1991), Panagopoulous and Spiliotis (1998), Nell (2001), Vera (2001), Lavoie (2005), Haghighat (2011)).

Industrial Production

Along with the gross domestic product (GDP) and the gross national product (GNP), industrial production is conventionally used in applied studies as a measure for the level of real economic activity, which is amongst the determinants of stock market prices. According to Maysami et al. (2004), an increase in industrial production signalizes a rise in economic growth. A positive relationship between industrial production and stock prices is hypothesized in the literature (Chen et al. (1986), Maysami et al. (2004), Rahman et al. (2009)), on the grounds that expected future cash flows are positively related to industrial production. Theoretically, a rise in the present value of the firm due to higher corporate earnings encourages investments in the stock market, ultimately leading to an increase in stock prices.

Money Supply

In spite of the amount of existing surveys and of the extensive inclusion of money supply as an explanatory factor for stock price levels, the impact of money supply on stock markets still remains an empirical question (Mukherjee and Naka (1995)). According to Fama (1981), an increase in money supply may be followed by a rise in inflation and in discount rates, which would lead to a decrease in stock prices. However, Mukherjee and Naka (1995) argue that the economic stimulus resulting from a growth in liquidity might neutralize this negative effect by increasing corporate future cash flows and consequently stock prices. On the other hand, according to portfolio theory, an increase in money supply may trigger investors to shift from a low-risk portfolio based on non-interest bearing monetary assets to a riskier portfolio based on financial assets such as stocks. In practice, some authors (Mukherjee and Naka (1995), Maysami et al. (2004) and Ratanapakorn and Sharma, (2007)) establish a positive relationship between money supply and stock prices, whereas others (Rahman et al. (2009) and Mahedi (2012)) have found evidence that the stock market is negatively influenced by money supply.

Interest Rates

Previous studies (Mukherjee and Naka (1995), Fama and Schwert (1977)) consistently determine a negative relationship between interest rates and stock prices, on the grounds of two different reasonings. In the first place, a decrease in interest rates reduces the cost of borrowing money and thus increases the investment capacity of firms, which is arguably an incentive for a rise in stock prices. Due to the enhanced corporate profits, investors will be willing to pay a higher price for the stock since they expect to receive higher dividends in the future. Conversely, an increase in interest rates is thought to increase the cost of trading on the stock market, since a large amount of stocks are purchased using borrowed money. As a consequence, investors will request that the rates of return be sufficiently high to compensate for the transaction costs, hence lowering demand and stock prices. The nominal contracting hypothesis postulated by French et al. (1983) establishes a link between the sensitivity of corporate stock returns to the term structure of the firm's holding of nominal assets and liabilities. As maintained by Fama (1975, 1976), Fama and Gibbons (1982) and Nelson and Schwert (1977), changes in expected and unexpected inflation are responsible for unanticipated movements in interest rate levels. Since the relationship between stock prices and inflation is positive when the firm's nominal asset holdings surpass the firm's nominal liabilities, while *ceteris paribus* it is negative when the opposite is true, then the stock market sensitivity to changes in interest rates is fundamentally linked to the firm's balance sheet composition (Flannery and James (1984)). Although the relationship between interest rates and stock returns is typically assumed to be negative, a few studies have found evidence of a positive linkage between short-term interest rates and stock prices (Asprem (1999), Mayasami and Koh (2000), Aspergis and Eleftherion (2002), Maysami et al. (2004)). As argued by Shiller and Beltratti (1992), this positive effect can be justified by the fact that a rise in interest rates may generate expectations about an increase in future fundamentals such as dividend payments, hence proving to be beneficial for corporate stock returns.

Crude Oil Price

The existing literature on the relationship between the price of crude oil and stock markets, which traces back to Hamilton (1983), advocates that, since crude oil is a fundamental input for production, an increase in the price of oil would have a negative impact on economic output, lowering economic activity in nearly all sectors. The resulting decrease in corporate future earnings is expected to negatively influence stock performance, causing a fall in stock prices. As argued bj Gjerde and Saettem (1999), however, this negative relationship is expected to hold for oil importing countries, while oil exporting countries are expected to profit from a rise in the price of oil.

Inflation Rate

A negative relationship between the rate of inflation and stock prices has been hypothesized by a substantial number of studies (Fama and Schwert (1977), Chen et al. (1986), Nelson (1976), Jaffe and Mandelker (1976), Mukherjee and Naka (1995), *inter alia*). A rise in inflation provokes a shift of resources from investment to consumption, reducing the volume of traded stocks and hence stock prices. As a general rule, monetary policies provide for economic tightening measures to counteract a rise in inflation, which causes the nominal risk-free rate and consequently discount rates to increase in a dividend-discount valuation model of type

$$\mathrm{P}_t = \frac{\mathrm{D}_{t+1}}{(k-g)}$$

where P is the stock price at present time t, $\mathrm{D}_{\mathrm{t+1}}$ are the dividends payed after the first period, g is the growth rates of dividends, assumed to be constant, and k is the rate of return of the stock. According to DeFina (1991), due to nominal contracts that prevent firm's costs and revenues to instantaneously adjust, cash flows and inflation grow at different paces. As a consequence, the negative effect resulting from a rise in the discount rate is not offset by the higher cash flows that follow inflation. However, Ratanapakorn and Sharma (2007) point out at a positive relationship between inflation and stock prices, whereas Gjerde and Saettem (1999) argue that the relationship is not significant.

The variables used in this survey are all expressed as natural logarithms, except for interest rates which is left in its original level-form, since it is a percentage figure. As can be seen in Figure A.1 in Appendix A, which displays the time path of the untransformed series, the variables DJX, IP, M1 and CPI exhibit a general trend of exponential growth. In this case, a logarithmic transformation appears reasonable, since the logarithm of a variable that grows at a

Variable	Definition
DJX	Natural logarithm of the price-valued weighted average of the month-end closing prices of 30 major American shares listed on the Dow Jones Averages[4] .
IP	Natural logarithm of the month-end U.S. industrial production index.
M1	Natural logarithm of the month-end U.S. narrow money supply (M1).
TB3M	The month-end U.S. 3-month Treasury bill secondary market rate.
OIL	Natural logarithm of the month-end U.S. West Texas Intermediate price for crude oil.
CPI	Natural logarithm of the month-end U.S. consumer price index.

Table 6.1: Description of the variables

Variable	Mean	Std Dev	Minimum	Maximum	Skewness	Kurtosis
DJX	8.96	0.58	7.72	9.81	-0.7	-0.82
IP	4.46	0.17	4.13	4.68	-0.73	-0.95
M1	7.19	0.35	6.65	8.03	0.8	-0.05
TB3M	3.15	2.44	0.01	8.82	0.17	-1.07
OIL	4.19	0.72	2.97	5.52	0.33	-1.38
CPI	5.19	0.19	4.8	5.47	-0.17	-1.15

Table 6.2: Descriptive statistics

[4]http://www.djaverages.com/.

constant rate will grow as a linear function of time (Hendry (1995))[3]. On the other hand, taking the logarithm of interest rates, which are expected to vary around a fixed level, appears superfluous. Moreover, since many economic time series have non-normal empirical distributions, the logarithmic transformation is often employed to achieve a model which provides a better fit to the classical linear model assumptions, as well as to reduce extreme values and diminish the effect of outliers in the data-set (Wooldridge (2003)). The macroeconomic variables and the stock index used in this survey are defined in Table 6.1, while Figure 6.1 shows the time path of the series. From visual inspection, the time series appear to be non-stationary and the upward trend of DJX, IP, M1 and CPI is clearly noticeable. A simultaneous drop in DJX, IP, TB3M and OIL can be seen around the year 2008-2009, which corresponds to the negative impact of the global financial crisis. Table 6.2 contains the descriptive statistics of the macroeconomic variables and the stock index. The values for the skewness and kurtosis indicate that none of the considered time series apparently follows a normal distribution. From the correlation coefficients displayed in Table 6.3, it appears that IP and OIL have the highest correlation with the Dow Jones index, while the correlation between TB3M and DJX is negative and has the lowest score.

	IP	M1	TB3M	OIL	CPI
DJX	0.9833481	0.7833289	-0.6547373	0.9184908	0.7074254

Table 6.3: Correlation between DJX and the macro-variables

[3] Constant percentage growth of a variable y implies that

$$\frac{\frac{dy}{dt}}{y} = g$$

Rearranging the terms and integrating both sided of the equation yields

$$\int \frac{dy}{y} = \int g\, dt$$

$$log(y) = gt + c$$

where c is a constant of integration and g is the percentage growth rate.

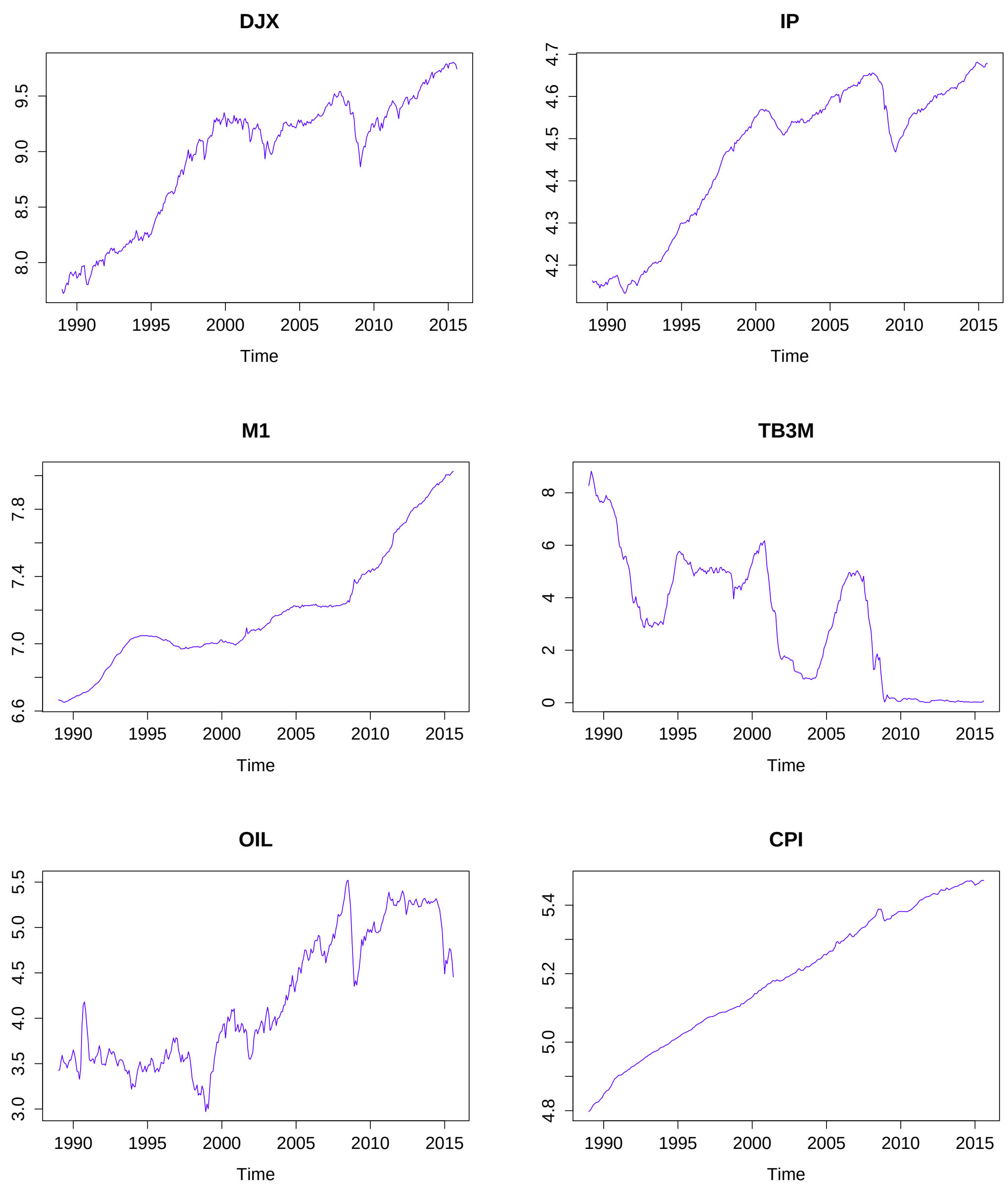


Figure 6.1: Time plot of all series

6.2 Unit Root Tests

6.2.1 Augmented Dickey-Fuller Test

In order to test whether the stock index and the macroeconomic variables contain a unit root, the augmented Dickey-Fuller (ADF) test is applied to each time series. Recall from Section 3.1 that the ADF test regression (3.3) can include either an intercept or a linear time trend. In this implementation of the ADF test, the following three variations of the test regression are considered:

$$\Delta y_t = c + Dt + \psi y_{t-1} + \sum_{i=1}^{p} \theta_i \Delta y_{t-i} + u_t \tag{6.1}$$

$$\Delta y_t = c + \psi y_{t-1} + \sum_{i=1}^{p} \theta_i \Delta y_{t-i} + u_t \tag{6.2}$$

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \theta_i \Delta y_{t-i} + u_t \tag{6.3}$$

The R function `ur.df()` from the package `urca` runs the ADF test. By setting the argument `type="trend"`, `type="drift"` or `type="none"`, it is possible to estimate test regression (6.1), (6.2) or (6.3), respectively. Additionally, the argument `selectlags = c("AIC", "BIC")` permits to use the Akaike or the Bayesian-Schwartz information criterion (Section 2.1) for selecting the optimal lag length. The command

```
ADF.tc <-ur.df(DJX, lags=10, type="trend", selectlags="BIC")
```

estimates regression (6.1) for the Dow Jones stock index. According to the BIC, serially uncorrelated errors are achieved by including one lagged endogenous variable in the test regression, as can be seen from the summary output in Table 6.4. The plot of the residual autocorrelations and partial autocorrelations in Figure 6.2 confirms this finding.

The results of the ADF test are obtained by comparing the test statistics with the appropriate critical values from Table 6.5. Performing a pure unit root test —that is, testing the null hypothesis $H_0 : \psi = 0$ against $H_1 : \psi \neq 0$, returns the τ_3 statistic, while testing the null hypotheses $\psi = c = 0$ and $\psi = Dt = 0$ by means of an *F*-type test, yields the test statistics ϕ_2 and ϕ_3, respectively. According to the test results, the null hypothesis $\psi = 0$ cannot be rejected for each of the time series at all significance levels, indicating that they are all unit root processes. The null hypothesis $\psi = c = 0$ cannot be rejected for M1 and CPI, which behave as random walks

Variable	Estimate	Std. Error	*t*-value	Pr(>|t|)
(Intercept)	0.1454	0.0806	1.804	0.0722
z.lag.1	-0.01680	0.009949	-1.688	0.0923
tt	0.000007045	0.00006079	1.159	0.2474
z.diff.lag	0.01258	0.05726	0.220	0.8263

Table 6.4: **ADF test**: Regression with constant and trend for DJX

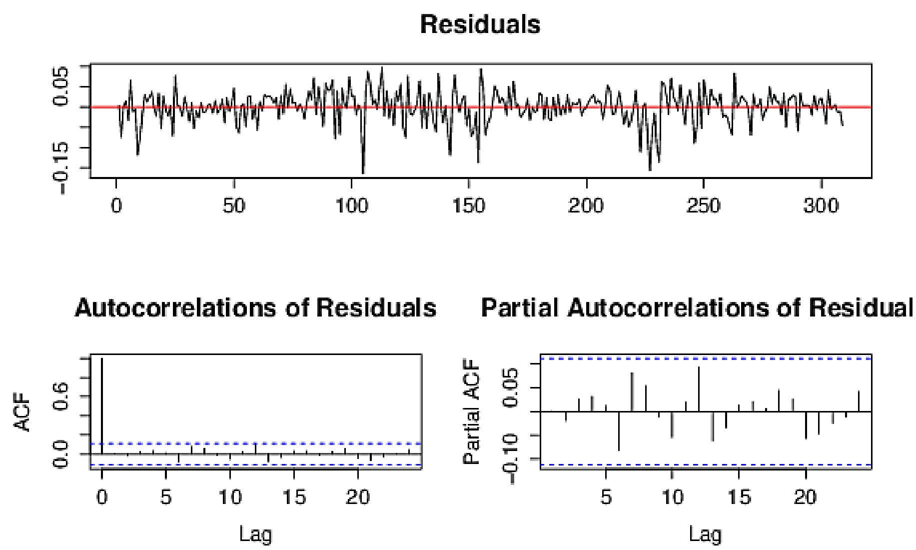


Figure 6.2: **ADF test**: Residual diagnostic for DJX

with drift, while the presence of a trend (hypothesis $\psi = Dt = 0$) is rejected for all series at the 5% significance level. The order of integration of the analyzed time series is determined by supplying their first difference as an argument for the ADF test: rejecting the null hypothesis of a unit root indicates that the series are I(1). As can be seen from the test results in Table 6.6, the first difference of each time series is stationary and hence the variables are I(1).

Test Statistics	τ_3	ϕ_2	ϕ_3
DJX	-1.6884	3.2461	1.7743
IP	-1.8999	2.4142	2.071
M1	0.1677	6.0176	1.6157
TB3M	-2.8475	3.0986	4.1564
OIL	-2.7535	2.6394	3.8815
CPI	-2.298	23.2502	6.1032
Critical Values	τ_3	ϕ_2	ϕ_3
1%	-3.98	6.15	8.34
5%	-3.42	4.71	6.30
10%	-3.13	4.05	5.36

Table 6.5: **ADF test**: Test statistics and critical values

Variable	Statistic	1%	5%	10%
Δ DJ	-12.2463	-2.58	-1.95	-1.62
Δ IP	-4.3845	-2.58	-1.95	-1.62
Δ M1	-2.9627	-2.58	-1.95	-1.62
Δ TB3M	-5.8045	-2.58	-1.95	-1.62
Δ OIL	-10.1554	-2.58	-1.95	-1.62
Δ CPI	-4.7092	-2.58	-1.95	-1.62

Table 6.6: **ADF test**: Order of integration

6.2.2 Phillips-Perron Test

The Phillips-Perron (PP) test, introduced in Section 3.2, is a unit root test similar to the ADF test, which however relieves the error process from the i.i.d. assumption. The test regression is allowed to contain both an intercept and a time trend, or an intercept only, as shown in equations (3.12) and (3.13). In this example, the R function `ur.pp()` from the package `urca` is used to run the PP test. The argument `use.lag` permits to select manually the number of lags to include for computing the long-run variance from equation (3.15). Otherwise, the lag length can automatically be set equal to $4(T/100)^{\frac{1}{4}}$ or $12(T/100)^{\frac{1}{4}}$ by choosing `lags="short"` or `lags="long"`. The test regressions (3.12) or (3.13) can be selected via the argument `model= c("constant", "trend")`. The command line

```
pp.test <- ur.pp(DJX, type="Z-tau", model="trend", lags="long")
```

runs the PP test for the DJX price series by estimating the test regression (3.12).

Test Statistics	$Z(\tau_\alpha)$	$Z(\tau_\mu)$	$Z(\tau_\beta)$
DJX	-1.8245	1.5861	1.2771
IP	-1.3879	-0.6584	1.0826
M1	0.0292	-1.4598	0.7667
TB3M	-2.6951	-1.6728	-1.9114
OIL	-2.0538	1.0069	1.6036
CPI	-2.3553	1.4309	2.1331
Critical values	$Z(\tau_\alpha)$	$Z(\tau_\mu)$	$Z(\tau_\beta)$
1%	-3.99	3.78	3.53
5%	-3.42	3.11	2.79
10%	-3.13	2.73	2.38

Table 6.7: **PP test**: Test statistics and critical values

The test statistics $Z(\tau_\alpha)$, $Z(\tau_\mu)$ and $Z(\tau_\beta)$ are retrieved with the method `summary`. The $Z(\tau_\alpha)$ statistics tests the null hypothesis that $\phi = 1$ against the alternative that $\phi \neq 1$, while $Z(\tau_\mu)$ tests the null hypothesis that $\mu = 0$ against the alternative that the true model contains an intercept. $Z(\tau_\beta)$ tests the null hypothesis that the true model is (3.13), i.e. $\lambda = 0$, against the alternative that (3.12) is the true data generating process. The critical values for the $Z(\tau_\alpha)$ statistics are stored in the slot `object@cval`, while the critical values for the $Z(\tau_\mu)$ and the

Variable	Test Statistic	1%	5%	10%
Δ DJX	-17.6772	-3.99	-3.43	-3.14
Δ IP	-16.8309	-3.99	-3.43	-3.14
Δ M1	-18.3587	-3.99	-3.43	-3.14
Δ TB3M	-12.9117	-3.99	-3.43	-3.14
Δ OIL	-11.8993	-3.99	-3.43	-3.14
Δ CPI	-10.8506	-3.99	-3.43	-3.14

Table 6.8: **PP test**: Order of integration

$Z(\tau_\beta)$ statistics can be found in Dickey and Fuller (1981). The values of the test statistics in Table 6.7 confirm the ADF test result of unit root nonstationarity for the considered time series, while the null hypotheses $\mu = 0$ and $\lambda = 0$ are not rejected for all series, which behave as pure random walks according to the Phillips-Perron test. The results of the PP test applied to the first difference of each variable, stored in Table 6.8, indicate that the time series are all I(1).

6.2.3 KPSS Test

The KPSS test, introduced in Section 3.3, is a Lagrange multiplier test used for testing the null hypothesis of trend or level-stationarity against the alternative hypothesis of a unit root. It is often employed as a countercheck for the ADF and PP test; if the results of the KPSS test are in contrast with those of unit root tests, the rule of thumb is to rely on the output from the stationarity test (Pfaff (2006)). The KPSS test is implemented in R with the function `ur.kpss()` from the package `urca`, which estimates the test regression 3.16. The argument `type` can be set equal to `mu` or `tau` to perform either a level-stationarity or a trend-stationarity test. In both cases ε_t in (3.16) is stationary under the null hypothesis, so that y_t is level-stationary when $\lambda = 0$ or trend-stationary when $\lambda \neq 0$. The Bartlett window parameter used for estimating the long-run variance (3.18) can be inputted either manually via the argument `use.lag` or selected automatically with `lags=c("short", "long")`, which produces the values $4(T/100)^{\frac{1}{4}}$ or $12(T/100)^{\frac{1}{4}}$. The command line

```
KPSS.mu <- ur.kpss(DJX, type = "mu", lags="short")
```

performs the level-stationary version of the KPSS test for the Dow Jones stock index. Since in Table 6.9 the values of the test statistics are bigger than the critical values, the null hypothesis of stationarity is rejected for each time series at the 5% significance level.

Test Statistics	DJX	IP	M1	TB3M	OIL	CPI
LM_μ	4.5477	4.5428	4.6162	3.6679	4.6844	5.4024
LM_τ	0.9578	1.0728	0.9222	0.1475	0.5965	0.6406
Critical Values	1%		5%		10%	
LM_μ	0.739		0.463		0.347	
LM_τ	0.216		0.146		0.119	

Table 6.9: **KPSS test**: Test statistics and critical values

6.2.4 Andrews-Zivot Structural Break Test

Since the presence of a structural break can often corrupt the results of DF-type tests (Perron (1989, 1990)), we run the unit root test developed by Andrews and Zivot (Section 3.4), which is robust with respect to endogenous structural shifts. The change is allowed to occur either in the levels, in the growth rate or in both the levels and the growth rate of a stochastic process. Models (3.21), (3.22) and (3.23) are specified accordingly to where the break occurs. Under the null hypothesis, the process y_t is a random walk with drift and no structural break occurs, while under the alternative, y_t is trend-stationary with a structural shift happening at time $\lambda = T_B/T$. The R function `ur.za()` from the contributed package `urca` implements the Andrews-Zivot test. Setting the functional argument `model` equal to `intercept`, `trend` or `both`, allows to choose between models (A), (B) and (C), respectively. The number of lagged endogenous variables to be included in the test regressions (3.21), (3.22) and (3.23) is determined via the argument `lags`. In R Code 6.1, Model (C) is estimated for the Dow Jones price series. The optimal lag length is determined by means of the AIC (command line 2) and the Breusch-Godfrey autocorrelation test from the package `lmtest` is used to verify that there is no residual serial correlation in the error process (command line 3).

```
1 za.test <- ur.za(DJX, model = "both", lag = 1)
2 AIC(eval(attributes(za.test)$testreg))
3 bgtest(attributes(za.test)$testreg, order=1)$p.value
```

R Code 6.1: Andrews-Zivot test for DJX

As can be seen from the graphs in Figure A.2 in Appendix A, none of the series has a structural break — which would otherwise be marked by a dashed vertical line, in either the levels or the growth rate. The unit root test results displayed in Table 6.10 confirm that all series are non-stationary without a structural break, at all significance levels.

Variable	Statistic	1%	5%	10%
DJX	-3.5605	-5.57	-5.08	-4.82
IP	-3.1806	-5.57	-5.08	-4.82
M1	-2.7801	-5.57	-5.08	-4.82
TB3M	-3.3783	-5.57	-5.08	-4.82
OIL	-3.7765	-5.57	-5.08	-4.82
CPI	-4.658	-5.57	-5.08	-4.82

Table 6.10: **Andrews-Zivot test**: Test statistics and critical values

6.2.5 HEGY Test for Seasonal Unit Roots

So far we have employed tests that are based on the assumption that the analyzed time series have only one root at the zero frequency. Since we are working with some series that are not seasonally adjusted —namely DJX, TB3M and OIL, it is possible that unit roots exist at frequencies other than zero, i.e. that these processes are seasonally integrated. Hylleberg et al. (1990) have developed a testing procedure that allows to test for the presence of seasonal unit roots, under the assumption that the DGP is an AR(p) with order p equal to the frequency of the data.

The HEGY test (Section 3.5) for monthly data is based on the test regression (3.35). Table 3.1 from Section 3.5 summarizes the null hypotheses and corresponding roots for the application of the HEGY test to monthly data.

The function `HEGY.test()` from the package `uroot` implements the HEGY test in R. Which deterministic terms should be added to the test regression is specified with the argument `itsd`, which is a three element vector. Setting `itsd=c(1,0,c(0))` includes an intercept in (3.35), `itsd=c(0,1,c(0))` adds a linear time trend while `itsd=c(0,1,c(1:11))` allows to include up to eleven seasonal dummy variables in the test regression. The number of lagged seasonal differences, used in (3.35) to whiten the error process, can be selected with the argument `selectlags`. The methods available for lag selection are the AIC, BIC, the Ljung-Box statistic or one can specify `"signf"` to maintain only the significant lags. The command line

```
hegy.test <- HEGY.test(wts = DJX, itsd = c(1, 0, c(1:11)),
+ selectlags = list(mode = "signf", pmax=NULL))
```

estimates a test regression with an intercept and eleven seasonal dummies for the Dow Jones stock index. The critical values for the HEGY test can be found in Franses and Hobijn (1997).

	Test Statistics			Critical Values		
Null Hypothesis	DJX	TB3M	OIL	1%	5%	10%
$\pi_1 = 0$	-1.719	-2.309	-1.235	-3.28	-2.76	-2.47
$\pi_2 = 0$	-5.508	-5.503	-5.039	-3.34	-2.76	-2.47
$\pi_3 = \pi_4 = 0$	20.730	24.656	34.895	8.35	6.27	5.28
$\pi_5 = \pi_6 = 0$	32.204	18.389	38.675	8.40	6.28	5.22
$\pi_7 = \pi_8 = 0$	29.388	33.775	29.556	8.32	6.21	5.21
$\pi_9 = \pi_{10} = 0$	25.891	41.811	26.109	8.34	6.22	5.23
$\pi_{11} = \pi_{12} = 0$	30.258	36.520	17.084	8.27	6.21	5.26

Table 6.11: **HEGY test**: Test statistics and critical values

The results of the test are stored in Table 6.11. Rejection of the null hypothesis $\pi_1 = 0$ indicates that the process has no unit root at the zero frequency, while rejection of the hypotheses

$\pi_i = 0$ for $i = 2, ..., 12$ indicates that the process has no unit roots at frequencies other than zero. The null hypothesis of a unit root at the zero frequency ($\pi_1 = 0$) cannot be rejected for all of the considered series, while the null hypothesis of seasonal unit roots $\pi_i = 0$ for $i = 2, ..., 12$ can be rejected at all significance levels, indicating that the series are not seasonally integrated.

6.3 Cointegration and Error Correction

6.3.1 Engle-Granger Test

In the previous sections, it was established that all of the considered series are I(1) processes and have neither structural breaks nor unit roots at seasonal frequencies. Now, the series are pairwise tested for cointegration by means of the Engle and Granger test (Section 6.3.1), which is aimed at detecting cointegration between the I(1) elements of a $(n \times 1)$ vector $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})'$ by estimating the long-run equilibrium equation

$$y_{1t} = \beta_2 y_{2t} + \beta_3 y_{3t} + ... + \beta_n y_{nt} + u_t \tag{6.4}$$

and then testing for the presence of a unit root in the residual process u_t by means of an ADF or PP test. The null hypothesis of no cointegration is rejected if u_t is found to be stationary. Following Engle and Granger's representation theorem (Section 4.4), if the hypothesis of a unit root is rejected for the residual process of equation (6.4), an error correction model (ECM) can be specified for the considered variables. For $n = 2$ the ECM has form

$$\begin{aligned} \Delta y_{1t} &= c_1 + \alpha_1 \hat{u}_{t-1} + \sum_{i=1}^{p-1} \gamma_{11}^i \Delta y_{1t-i} + \sum_{i=1}^{p-1} \gamma_{12}^i \Delta y_{2t-i} + u_{1t} \\ \Delta y_{2t} &= c_2 + \alpha_2 \hat{u}_{t-1} + \sum_{i=1}^{p-1} \gamma_{21}^i \Delta y_{2t-i} + \sum_{i=1}^{p-1} \gamma_{22}^i \Delta y_{1t-i} + u_{2t} \end{aligned} \tag{6.5}$$

where the error correction term $\hat{u}_{t-1} = y_{1t-1} - \hat{\beta}_2 y_{2t-1}$ is the estimate of the lagged error from (6.4) for $n = 2$ and α_i $(i = 1, 2)$ is the speed of adjustment coefficient, which measures the proportion of disequilibrium that is corrected at time t. In order to prevent the divergence of the system from its long-run equilibrium, the sign of the coefficient α is required to be negative (Hamilton (1994)). In R Code 6.2, the long-run equation

$$\mathrm{DJX}_t = c_1 + \beta_2 \mathrm{IP}_t + u_{1t}$$

is estimated by OLS and the regression residuals are stored as an object `error`. Table 6.12 displays the results of the ADF test on the residuals obtained from regressing DJX on the single macroeconomic variables. In this application, the critical values for the ADF test are provided by Engle and Yoo (1987) or MacKinnon (1991), since the test is run on the estimated error term $\hat{u}_t$ from a spurious regression. The values of the test statistics imply that only the Dow Jones index and the industrial production series are cointegrated.

```
eq.1 <- lm(DJX ~ IP)
error.1 <- ts(resid(eq.1), start=c(1989, 1), end=c(2015, 8),
    frequency=12)
```

R Code 6.2: Engle-Granger Test

Variable	ADF	1%	5%	10%
DJX - IP	-3.8345	-3.78	-3.25	-2.98
DJX - M1	-1.2239	-3.78	-3.25	-2.98
DJX - TB3M	-1.723	-3.78	-3.25	-2.98
DJX - OIL	-1.7992	-3.78	-3.25	-2.98
DJX - CPI	-1.6454	-3.78	-3.25	-2.98

Table 6.12: **Engle-Granger Test**: Pairwise regression

Although the Engle-Granger test did only detect cointegration between DJX and IP when applied pairwise, it is still possible that the stock index is cointegrated with the set of macroeconomic factors. Hence, a long-run equilibrium equation of type (6.4) is estimated for the entire data-set

$$\mathrm{DJX}_t = c + \beta_2 \mathrm{IP}_t + \beta_3 \mathrm{M1}_t + \beta_4 \mathrm{TB3M}_t + \beta_5 \mathrm{OIL}_t + \beta_6 \mathrm{CPI}_t + u_t \tag{6.6}$$

and both and ADF and a PP test are run on the residuals of this equation. In R Code 6.3, the series entering the long-run equation are stored as an object `dataset` and the function `window` is used to select the sample period (command line 1). Next, equation (6.6) is estimated by OLS and the regression residual is stored as `error.DJX` (command lines 2 and 3). The result of the ADF and PP test on the residuals from equation (6.6), displayed in Table 6.13, indicate the presence of a cointegrating vector in the system at the 5% significance level[5].

Test	Statistic	1%	5%	10%
ADF	-4.8444	-4.99	-4.40	-4.14
PP	-4.9137	-4.99	-4.40	-4.14

Table 6.13: **Engle-Granger Test**: Regression of DJX on macro-variables

Since the system of variables is cointegrated, an ECM of the the form of equation (6.5) is specified with DJX as the dependent variable. In R Code 6.3, the function `embed` is used to create a matrix whose elements are the first difference and the lagged first difference of the considered variables, and the respective column names are assigned to each element (command lines 4 and 5). Finally, the regression residuals are lagged by one period and the ECM equation is estimated (command lines 6 and 7). Table 6.14 displays the summary output of the ECM. The error correction term is highly significant and has the correct sign; its estimated value indicates that approximately 12% of the previous period disequilibrium was corrected. The lagged first difference of IP and TB3M enter significantly into the ECM equation, whereas the other regressors do not respond significantly to a change in DJX. When cointegration is detected between two variables, there should be Granger causality (Section 4.3) in at least one direction (Enders (1995)). Due to the significant speed of adjustment coefficient and lagged differences $\Delta\mathrm{IP}_{t-1}$ and $\Delta\mathrm{TB3M}_{t-1}$, industrial production and interest rates are said to Granger cause the U.S. stock market index.

[5]The critical values for the ADF and PP test are retrieved from Phillips and Ouliaris (1990).

```r
dataset <- window(cbind(DJX, IP, M1, TB3M, OIL, CPI), start=c(1989, 1), end=c(2015, 8))
eq.DJX <- lm(DJX ~ IP + M1 + TB3M + OIL + CPI, data=dataset)
error.DJX <- ts(resid(eq.DJX), start=c(1989, 1), end=c(2015, 8), frequency=12)
dataset2 <- ts(embed(diff(dataset), dim=2), start=c(1989, 1), end=c(2015, 8), freq=12)
colnames(dataset2) <- c("DJX.d", "IP.d", "M1.d", "TB3M.d", "OIL.d", "CPI.d", "DJ.d1", "IP.d1", "M1.d1", "TB3M.d1", "OIL.d1" "CPI.d1")
error.ecm <- lag(error.DJX, k=-1)
ecm.eq <- lm(DJX.d ~ error.ecm + DJX.d1 + IP.d1 + M1.d1 + TB3M.d1 + OIL.d1 + CPI.d1, data=dataset2)
```

R Code 6.3: ECM for the stock index

Variable	Estimate	Std. Error	*t*-value	Pr(>|t|)
(Intercept)	0.002496	0.003428	0.728	0.4671
ECT	-0.118879	0.027582	-4.310	0.0000219
ΔDJX_{t-1}	-0.039561	0.054681	-0.724	0.4699
ΔIP_{t-1}	1.817178	0.382837	4.747	0.00000316
$\Delta M1_{t-1}$	-0.314277	0.264501	-1.188	0.2357
$\Delta TB3M_{t-1}$	-0.029692	0.012464	-2.382	0.0178
ΔOIL_{t-1}	-0.032450	0.031055	-1.045	0.2969
ΔCPI_{t-1}	0.993127	0.969465	1.024	0.3064

Table 6.14: **Engle-Granger Test**: ECM for the stock index

6.3.2 Johansen Test

Although the result of the Engle-Granger test indicated that the Dow Jones index is cointegrated with the set of macroeconomic variables, it is possible that more than one cointegrating relationship exists. In fact, in the instance that a time series vector $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})$ has $n > 2$ elements, there may exist up to $n - 1$ independent linear combinations of the elements of $\boldsymbol{Y}_t$ whose order of integration is lower than that of the original series (Lütkepohl, (2006)). The drawback of the Engle-Granger procedure is that, even in the case of $n > 2$, only a single cointegrating relationship is found, although the system may contain up to $n-1$ distinct cointegrating relationships. The Johansen cointegration test (Section 5.2) allows to determine the number of cointegrating vectors in a system with $n > 2$ variables by estimating an vector error correction model (VECM) such as (5.8) from Section 4.4, which is allowed to contain deterministic terms such as a vector of constants or seasonal dummy variables. Johansen's procedure focuses on the determination of the rank of matrix $\boldsymbol{\Pi}$ from equation (5.8), which must fulfill the condition $0 < \text{rank}(\boldsymbol{\Pi}) = h < n$ in order for the $(n \times h)$ matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to exist such that $\boldsymbol{\alpha\beta}' = \boldsymbol{\Pi}$ and $\boldsymbol{\alpha\beta'Y}_t \sim I(0)$. The rank of $\boldsymbol{\Pi}$ is equal to the number of cointegrating vectors in the system, which are contained in $\boldsymbol{\beta}$, while $\boldsymbol{\alpha}$ measures the speed of the error correction mechanism. The trace statistic λ_{trace} from equation (5.9) tests the null hypothesis that there are at least h cointegrating vectors in $\boldsymbol{Y}_t$, whereas the maximum eigenvalue statistic $\lambda_{\max}$ from equation (5.10) tests the null hypothesis that the number of cointegrating vectors is h against the alternative that it is $h + 1$.

Lag	1	2	3	4	5	6	7	8	9	10	11	12
AIC	-46.83	-47.52	-47.54	-47.61	-47.66	-47.58	-47.60	-47.50	-47.44	-47.38	-47.31	-47.22

Table 6.15: VAR order selection by AIC

The R function `ca.jo()` from the package `urca` implements Johansen's procedure for determining the cointegration rank of $\boldsymbol{Y}_t = (\text{DJ}, \text{IP}, \text{M1}, \text{TB3M}, \text{OIL}, \text{CPI})$. The argument `type` can be set equal to `trace` or `eigen` to select between the trace and the maximum eigenvalue statistic, and the argument `transitory` specifies that the levels of $\boldsymbol{Y}_t$ enter the VECM at lag $t - 1$. The optimal number of lags to be included in the VECM equation is selected with the function `VARselect()` from the package `vars`, which returned a lag length of 12 according to the multivariate AIC (Section 4.3), as reported in Table 6.15. As can be seen from Table 6.16 and Table 6.17, the trace statistic rejects the null hypothesis of $h \leq 1$ in favor of $h = 2$ at the 5% significance level, whereas the maximum eigenvalue statistic detects the presence of one cointegrating vector in the system.

Hypothesis	Statistic	1%	5%	10%
$h \leq 5$	7.09	6.50	8.18	11.65
$h \leq 4$	17.18	15.66	17.95	23.52
$h \leq 3$	28.09	28.71	31.52	37.22
$h \leq 2$	47.46	45.23	48.28	55.43
$h \leq 1$	78.13	66.49	70.60	78.87
$h = 0$	128.76	85.18	90.39	104.20

Table 6.16: **Johansen test**: Trace test

Hypothesis	Statistic	1%	5%	10%
$h \leq 5$	7.09	6.50	8.18	11.65
$h \leq 4$	10.09	12.91	14.90	19.19
$h \leq 3$	10.91	18.90	21.07	25.75
$h \leq 2$	19.37	24.78	27.14	32.14
$h \leq 1$	30.66	30.84	33.32	38.78
$h = 0$	50.63	36.25	39.43	44.59

Table 6.17: **Johansen test**: Maximum eigenvalue test

Next, the R function `cajorls()` from the package `urca` is used to estimate a VECM such as (5.8) under the rank restriction $h = 1$. According to the VAR(12) order specification by AIC from the previous step, twelve lagged differences of the variables are included in the model. Be-

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
$\Delta \mathrm{IP}_{t-1}$	1.146551	0.480790	2.385	0.017869
$\Delta \mathrm{TB3M}_{t-1}$	-0.055165	0.016249	-3.395	0.000803
$\Delta \mathrm{OIL}_{t-1}$	-0.081119	0.036034	-2.251	0.025278
$\Delta \mathrm{CPI}_{t-2}$	-2.4483594	1.4659072	-1.670	0.096183
$\Delta \mathrm{CPI}_{t-4}$	-3.192945	1.489003	-2.144	0.033010
$\Delta \mathrm{M1}_{t-5}$	1.126398	0.327364	3.441	0.000684
$\Delta \mathrm{TB3M}_{t-6}$	-0.053017	0.017656	-3.003	0.002958
$\Delta \mathrm{CPI}_{t-6}$	-2.748904	1.467395	-1.873	0.062239
$\Delta \mathrm{OIL}_{t-8}$	0.085802	0.040807	2.103	0.036541
$\Delta \mathrm{CPI}_{t-8}$	-4.749164	1.474509	-3.221	0.001455
$\Delta \mathrm{OIL}_{t-9}$	0.087121	0.040976	2.126	0.034511
$\Delta \mathrm{IP}_{t-10}$	-0.947484	0.471873	-2.008	0.045772
$\Delta \mathrm{OIL}_{t-11}$	0.088659	0.041491	2.137	0.033625
$\Delta \mathrm{CPI}_{t-11}$	-2.188399	1.267127	-1.727	0.085444

Table 6.18: **VECM**: Significant regressors in the VECM equation for DJX

sides the cointegratig rank, `reg.number=1` is imposed as an argument so that only the VECM equation for DJX is shown. Table B.1 in Appendix B displays the estimated VECM equation for DJX, while Table 6.18 reports only the significant lags of each regressor. As can be seen, the lagged differences of each of the selected macroeconomic variables enter significantly in the VECM equation for the stock index. The industrial production coefficient is significant at lags 1 and 10 and has a negative sign at lag 10, which is in contrast to what is assumed in the literature (see Section 6.1). The coefficient of money supply enters the VECM at lag 5 with a positive sign, in accordance with the findings of Mukherjee and Naka (1995), Maysami et al. (2004) and Ratanapakorn and Sharma (2007), among others. The first difference of interest rates is significant at lags 1 and 6 and is consistently negative, as predicted by Mukherjee and Naka (1995) and Fama and Schwert (1977). Crude oil price enters the VECM at lags 1, 8, 9 and 11 and has a negative sign only in the first lag, in accordance with what is assumed by Gjerde and Saettem (1999). As hypothesized in the literature, inflation rate has a negative effect on the stock price at each significant lag.

```
dataset <- cbind(DJX, IP, M1, TB3M, OIL, CPI)
H1.trace <- ca.jo(dataset, type="trace", ecdet="none", K=12, spec="transitory")
vecm <- cajorls(H1.trace, r = 1)
beta <- H1.trace@V[,1]
alpha <- H1.trace@W[,1]
resids <- resid(vecm$rlm)
N <- nrow(resids)
sigma <- crossprod(resids)/N
beta.se <- sqrt(diag(kronecker(solve(crossprod(H1.trace@RK[, -1])), solve(t(alpha)%*% solve(sigma)%*%alpha))))
beta.t <- c(NA, beta[-1]/beta.se)
alpha.se <- sqrt(solve(crossprod(cbind(H1.trace@ZK%*%beta, H1.trace@Z1)))[1, 1]*diag(sigma))
alpha.t <- alpha/alpha.se
```

R Code 6.4: Estimated VECM with $h = 1$

	DJX_{t-1}	IP_{t-1}	$\mathrm{M1}_{t-1}$	$\mathrm{TB3M}_{t-1}$	OIL_{t-1}	CPI_{t-1}
$\boldsymbol{\beta}$	1.000000	-3.562367 $(-10.759460)^{***}$	-0.497635 $(-4.455171)^{***}$	-0.063711 $(-4.419530)^{***}$	-0.025205 (-0.319059)	1.018790 (1.454542)
$\boldsymbol{\alpha}$	-0.125331 $(-4.3792969)^{***}$	0.001301 (0.3250324)	0.021102 $(3.8071084)^{***}$	0.034419 (0.2972556)	-0.098349 (-1.6574874)	-0.005619 $(-3.6395815)^{***}$

Note: An '***' denotes statistical significance at the 0.1 percent. T-values are in parenthesis.

Table 6.19: **VECM**: Normalized eigenvector, weights and t-statistics

Table 6.19 displays the estimated cointegration matrix $\boldsymbol{\beta}$ and the loading matrix $\boldsymbol{\alpha}$ with their respective t-statistics in parenthesis, which are computed according to the procedure shown in R Code 6.4. Since the speed of adjustment coefficient $\boldsymbol{\alpha}$ is significant in the VECM equation for DJX, the stock index responds to deviations from long-run equilibrium and the proportion of disequilibrium that is corrected within one time period is 13%. The cointegrating vector is

$$\boldsymbol{\beta} = (1.00,\ -3.56,\ -0.50,\ -0.06,\ -0.03,\ 1.02)'$$

which can alternatively be expressed as

$$\mathrm{DJX} = 3.56\mathrm{IP} + 0.50\mathrm{M1} + 0.06\mathrm{TB3M} + 0.03\mathrm{OIL} - 1.02\mathrm{CPI} \tag{6.7}$$

Due to the logarithmic transformation, the values of the coefficients of DJX, IP, M1, TB3M , OIL and CPI in the cointegrating relationship can be regarded as long-term elasticities. The values of the t-statistics reported in Table 6.19 indicate that the long-run coefficients of IP, M1 and TB3M are statistically significant, whereas the coefficients of OIL and CPI are not significantly different from zero. A LR test by Johansen (1988) is performed on the cointegrating vector in order to determine if the coefficients in $\boldsymbol{\beta}$ contribute to the long-run equilibrium relationship. The test is implemented via the R function `blrtest()` from the package `urca`, which tests the validity of placing $n-m$ linear restrictions on the eigenvectors in $\boldsymbol{\beta}$ by estimating a restricted VECM with n variables and m restrictions on the cointegrating space. The null hypothesis of the LR test is defined as

$$H_0 : \boldsymbol{R}'\boldsymbol{\beta} = 0 \text{ or } \boldsymbol{\beta} = \boldsymbol{H}\boldsymbol{\phi} \tag{6.8}$$

where $\boldsymbol{H} = \boldsymbol{R}_\perp$ and $\boldsymbol{R}_\perp$ is the orthogonal complement of $\boldsymbol{R}$ such that $\boldsymbol{R}'\boldsymbol{R}_\perp = 0$ and $\boldsymbol{R}'\boldsymbol{H} = 0$. In (6.8), the restrictions placed on $\boldsymbol{\beta}$ are defined in the known $(n \times m)$ matrix $\boldsymbol{H}$, while the unknown $(m \times h)$ matrix $\boldsymbol{\phi}$ contains the reduced parameters that result from imposing m restrictions, where $h \leq m \leq n$. The null hypothesis H_0 is tested against the alternative hypothesis $H(h) : \boldsymbol{\pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ of h unrestricted cointegrating vectors by means of the λ_{res} test statistic from equation (5.12). Testing the significance of each element in $\boldsymbol{\beta}$ is equivalent to considering the set of hypotheses

$$H_{0,1} : \boldsymbol{\beta}_1 = 0 \text{ vs } H_{1,1} : \boldsymbol{\beta}_1 \neq 0$$

$$H_{0,2} : \boldsymbol{\beta}_2 = 0 \text{ vs } H_{1,2} : \boldsymbol{\beta}_2 \neq 0$$

$$H_{0,3} : \boldsymbol{\beta}_3 = 0 \text{ vs } H_{1,3} : \boldsymbol{\beta}_3 \neq 0$$

$$H_{0,4} : \boldsymbol{\beta}_4 = 0 \text{ vs } H_{1,4} : \boldsymbol{\beta}_4 \neq 0$$

$$H_{0,5} : \boldsymbol{\beta}_5 = 0 \text{ vs } H_{1,5} : \boldsymbol{\beta}_5 \neq 0$$

$$H_{0,6} : \boldsymbol{\beta}_6 = 0 \text{ vs } H_{1,6} : \boldsymbol{\beta}_6 \neq 0$$

and requires the specification of six different matrices of restrictions $\boldsymbol{H}_{0,j}$, $j = 1, ..., 6,$, of the form

$$\boldsymbol{H}_{0,1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{H}_{0,2} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{H}_{0,3} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{H}_{0,4} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{H}_{0,5} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{H}_{0,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Put differently, assume for example that we wish to test the hypothesis that the long-run coefficient of IP is zero. This is equivalent to testing the null hypothesis $H_{0,2}$, which can alternatively be expressed in terms of the cointegrating relationship as

$$\boldsymbol{\beta} = (\beta_1,\ 0,\ \beta_3,\ \beta_4,\ \beta_5,\ \beta_6)'$$

without placing any restrictions on the other long-term coefficients. Following the parametrization of (6.8), the matrix of restrictions $\boldsymbol{H}$ is constructed as $\boldsymbol{H}_{0,2}$ and the matrix of unknown parameters becomes $\boldsymbol{\phi} = (\beta_1,\ \beta_3,\ \beta_4,\ \beta_5,\ \beta_6)'$, such that the constraint $\boldsymbol{\beta} = \boldsymbol{H}\boldsymbol{\phi}$ is satisfied.

Variable	Statistic	p-value
DJX	14.79	0.000120
IP	17.7	0.000026
M1	2.8509	0.001142
TB3M	5.1903	0.008077
OIL	1.856	0.841478
CPI	6.6493	0.338327

Table 6.20: LR test of linear restrictions in $\boldsymbol{\beta}$

The resulting p-values in Table 6.20 indicate that the restrictions placed on $\boldsymbol{\beta}$ are rejected for IP, M1 and TB3M, whereas they are not rejected for OIL and CPI, suggesting that the coefficients of oil price and inflation are not significantly different from zero in the cointegrating vector. Additionally, the validity of the VECM is established given that the coefficient of DJX significantly contributes to the cointegrating relationship (Maysami and Koh (2000)). A test of the null hypothesis that the long-term relationship is given by

$$\boldsymbol{\beta} = (\beta_1,\ \beta_2,\ \beta_3,\ \beta_4,\ 0,\ 0)'$$

entails specifying the restriction matrix

$$\boldsymbol{H}_{0,7} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and yields a test statistic of 4.58 and a p-value of 0.10, confirming the assumption that OIL and CPI do not participate to the error correction mechanism. Similar results were obtained by Chen et al. (1986), who found that the oil price does not influence U.S. stock prices, and by Mayasami and Koh (2000) and Humpe and Macmillan (2007), who failed to establish the existence of a long-term relationship between inflation and the stock exchange of Singapore and Japan, respectively. In accordance to what is hypothesized in the literature (see Section 6.1), the long-run coefficient of industrial production in (6.7) has a positive sign. In our empirical results, money supply has a positive influence on stock returns, consistently with the findings of Mukherjee and Naka(1995), Maysami et al. (2004) and Ratanapakorn and Sharma (2007). The positive sign of the interest rates coefficient can be explained by the possibility that an increase in interest rates is followed by the anticipation of higher dividend returns (Shiller and Beltratti (1990)) and is in accordance with the empirical results of Asprem (1999), Mayasami and Koh (2000) and Aspergis and Eleftherion (2002). Although the coefficients of oil price is not significant, its positive sign confirms the theoretical findings of Gjerde and Saettem (1999), who maintain that an increase in oil price positively influences stock returns in oil exporting countries. Similarly to the results of Fama and Schwert (1977), Chen et al. (1986), Nelson (1976), Jaffe and Mandelker (1976) and Mukherjee and Naka (1995), the long-run relationship between stock prices and inflation is negative, albeit not significant.

6.4 Diagnostic Tests

Once the VECM model has been estimated, diagnostic tests can be performed in order to check whether the model adequately describes the relevant features of the process generating the analyzed data-set. Many of these tests focus on the residuals of the single or joint VECM equations and aim at ensuring that the residual process underlies the required assumptions. Formally, it is assumed that the error process $\boldsymbol{u}_t$ from a VECM such as (5.8) is $i.i.d.$ and normally distributed with mean 0 and covariance matrix $\boldsymbol{\Sigma}$. In order to conduct diagnostic tests on the residuals of a VECM in R, it is necessary to express the estimated VECM as a VAR model in levels. In Section 4.4, the procedure of deriving a VECM from a VAR model in levels has been shown; hence, it is possible to retransform a VECM in its level-VAR representation. The R function `vec2var()` from the package `vars` achieves this transformation, provided that the cointegration rank h of the VECM is supplied as an argument. Table B.2 in Appendix B displays the estimated coefficients of the VECM for $\boldsymbol{Y}_t = (\mathrm{DJX}, \mathrm{IP}, \mathrm{M1}, \mathrm{TB3M}, \mathrm{OIL}, \mathrm{CPI})$, expressed in its level-VAR form.

Variable	Mean	Std Dev	Minimum	Maximum	Skewness	Kurtosis
DJX	0.000000	0.033791	-0.145236	0.094666	-0.33	0.39
IP	0.000000	0.004728	-0.030099	0.015263	-0.74	5.08
M1	0.000000	0.006544	-0.034191	0.032393	0.11	4.28
TB3M	0.000000	0.136718	-0.480816	0.698258	0.08	3.60
OIL	0.000000	0.070060	-0.234624	0.358915	0.24	1.97
CPI	0.000000	0.001823	-0.006024	0.007794	0.19	1.58

Table 6.21: Descriptive statistics of VAR residuals

As a preliminary check of the VAR model adequacy in representing the DGP of the underlying data-set, the descriptive statistics of the single residual series are calculated and the residual plots, empirical densities and autocorrelograms are visually inspected. According to simulation studies, statistical inference in relation to VAR models is especially sensitive with respect to validating the assumptions of constant parameters, serially uncorrelated residuals and zero residual skewness, while it quite robust with respect to heteroskedastic residuals and excess kurtosis (Juselius and Hendry (2001)). As it is displayed in Table 6.21, the mean is not significantly different from zero for any of the considered time series. Although the residual distributions are quite symmetrical, the skewness and excess kurtosis do not correspond to a normal value of 0. From inspecting the residual plots of the time series, stored in Appendix A, there appears to be some significant outlier observations. One large outlier in the residuals of IP appears around 2007, the year of the outbreak of the global financial crisis. Money supply has one outlier in year 2002, TB3M has two long spikes in 1998 and 2008, OIL has one outlier in 1989 and CPI seems to have more than one between 2005 and 2010. The residual volatility of DJX and IP appears to be rather constant in time, but money supply has one period of higher variance between 2005 and 2010, and the volatility of the residuals of TB3M is very dissimilar in various parts of the sample period, indicating that the series is likely to be heteroskedastic. The presence of some ARCH effects in the residual series of OIL and CPI is not excluded, albeit this cannot be established by graphical inspection alone. The assumption of serially uncorrelated error terms is investigated by inspecting the autocorrelations and partial autocorrelations of the residual series. When there is no serial correlation in the residual series, most of the autocorrelations and partial autocorrelations will lie within an interval of width $\pm 2/\sqrt{T}$, where T is the number of observations, which corresponds to a 95% confidence interval. Although few exceptions are permitted, values that exceed $\pm 2/\sqrt{T}$ in the first lags are considered particularly suspicious and usually indicate model inadequacy (Lütkepohl and Krätzig (2004)). Although the small values of the autocorrelation coefficients are consistent with a white noise process for all residual series, the autocorrelations of the squared residuals clearly trespass the confidence band, which possibly signalizes a dependence structure in the second order moments of the residual distribution. Serial correlation is especially significant in the squared residuals of the interest rate series, giving rise to concern about the presence of volatility clusterings.

6.4.1 Jarque-Bera Normality Test

Although financial time series frequently do not to comply with Gaussian assumptions due to outliers, skewness, kurtosis and volatility clusters, many statistical techniques employed in cointegration analysis rely on the assumption of normality. For example, Johansen's likelihood ratio tests for determining the cointegration rank and Johansen's maximum likelihood (ML) estimator for VECM coefficients (Section 5.2) are derived under the assumption that the innovations follow an $i.i.d.$ normal distribution. Nevertheless, there is evidence that both techniques are still valid under non-normality. The impact of a non-Gaussian cointegration error distribution on the size and power of Johansen's cointegration rank test was investigated by Cheung and Lai (1993) and Gonzalo (1994). They found that the likelihood based trace and maximum eigenvalue statistics are robust with respect to skewness and kurtosis in the VECM residuals. In addition, Gonzalo (1994) provides Monte Carlo evidence that the ML estimate of a cointegrating relationship outperforms single equation methods and multivariate methods[6], in addition to being robust to non-Gaussian innovations. This is due to the fact that Johansen's ML estimator can be derived as a special case of a reduced rank simultaneous least squares estimator,

[6]Besides full information ML, the other estimation methods examined in Gonzalo's paper are ordinary least squares (Engle and Granger (1987)), nonlinear least squares (Stock (1987)), principal components (Stock and Watson (1988)) and canonical correlations (Bossaerts (1988)).

which underlies no specific assumptions about the distribution of the disturbance terms (*ibid.*).

Variable	Statistic	p-value
DJX	8.0067	0.01825
IP	366.14	2.2×10^{-16}
M1	241.29	2.2×10^{-16}
TB3M	170.39	2.2×10^{-16}
OIL	54.503	1.5×10^{-12}
CPI	34.96	2.6×10^{-8}
Multivariate	826	2.2×10^{-16}
Skewness	41.203	2.6×10^{-7}
Kurtosis	784.8	2.2×10^{-16}

Table 6.22: Univariate and multivariate Jarque-Bera test

In this section, the residuals of each individual VECM equation and the residual matrix of the VECM are tested for normality by means of the univariate and multivariate Jarque-Bera test from Section 2.3.1. The R function `normality.test()` from the package `vars` implements both versions of the Jarque-Bera test. When the argument `multivariate.only` is set equal to `FALSE`, the test statistics for the single residual series are computed along with the VECM residual matrix. The results of the univariate and multivariate Jarque-Bera test statistics are stored in Table 6.22. Due to the extremely low p-values, the null hypothesis that the residuals of the VECM and of the single VECM equations follow a normal distribution must be rejected. Also the test statistics for the skewness and kurtosis of the residual distribution are not in compliance with the third and fourth moment vectors of a normally distributed variable. This results are not surprising considering the values for the skewness and kurtosis from the descriptive statistics in Table 6.21.

6.4.2 Heteroskedasticity Test

The presence of conditional heteroskedasticity in economic and financial time series has received much attention in the literature ever since the introduction of the ARCH model by Engle (1982). It is by now well-documented that inhomogeneous variance can be the cause of invalidity of standard asymptotic tests in time series models [7]. In practice, is quite common that the residual process of financial time series like interest rates exhibit ARCH effects, and their presence does not necessarily indicate that the model is inadequate for describing the underlying DGP (Lütkepohl (2004)). Although Cavaliere et al. (2010a, 2010b) establish the validity of the likelihood ratio (LR) tests proposed by Johansen (1996) even in the presence of conditional heteroskedasticity in the residual series, they find that the tests are oversized in small and finite samples, i.e. they tend to overreject the null hypothesis, which may lead to finding too many cointegrating relationships. Also the empirical size of the Dickey-Fuller test tends to be higher than the nominal size in the presence of ARCH errors, and the null hypothesis of no cointegration is more frequently rejected (Kim and Schmidt (1993)). Furthermore, standard model selection criteria (Section 4.3) may be biased towards selecting a large number of lags when the variables exhibit ARCH effects (Catani (2013)). However, Lee and Tse (1996) examined the size and power of Johansens's LR test when the cointegration residuals were modeled by a GARCH(1,1) with student-*t* and normal distribution, and found that, in spite of its tendency to overrejection, the test performance is still higher with respect to other cointegration tests such as the Dickey-Fuller (DF) test and the cointegration regression Durbin-Watson (CRDW) test. Moreover, Cheung and Lai (1993) and Gonzalo (1993) report that Johansen's maximum likelihood estimation method of the cointegrating vector is robust with respect to non-normality, overparametrization as well as heteroskedasticity. Monte Carlo simulation studies demonstrate that the performance of the ML estimator remains superior to that of other methods even when the standard assumptions for VECM disturbances are violated (*ibid.*).

The multivariate ARCH-LM test for residual heteroskedasticity (Section 2.3.2) is a Lagrange multiplier (LM) test used to test the null hypothesis of no ARCH effects in the residual matrix

[7] See for example, Pantula (1988), Haldrup (1994), Kim et al. (2002), Busetti and Taylor (2003) and Cavaliere and Taylor (2006, 2007).

Variable	Statistic	p-value
DJX	15.175	0.232
IP	3.5526	0.9902
M1	20.487	0.05842
TB3M	54.754	2.0×10^{-7}
OIL	26.399	0.009422
CPI	20.875	0.05224
Multivariate	5407.9	0.1303

Table 6.23: Univariate and multivariate ARCH-LM test

of a VECM model at lags ranging from 1 to q,

$$H_0 : \boldsymbol{B}_1 = ... = \boldsymbol{B}_q = 0$$

against the alternative of significant residual heteroskedasticity,

$$H_1 : \boldsymbol{B}_i \neq 0 \; (i = 1, ..., q)$$

based on the auxiliary test regression (2.9). The R function `arch.test()` from the package `vars` computes the univariate and multivariate ARCH-LM test statistics. Setting the argument `multivariate.only` equal to `FALSE` returns the results of both the univariate and multivariate ARCH-LM test. The number of lags to be included in the auxiliary regressions (2.8) and (2.9) is set to 12 via the functional arguments `lags.single` and `lags.multi`. Table 6.23 contains the results of the ARCH-LM test applied on the single residual series and on the residual matrix of the VAR-level model. As it was expected, there are no significant ARCH effects in the residuals of DJX, IP and M1 and the null hypothesis of homoskedasticity is not rejected for the residual matrix of the VECM. However, the residual series of TB3M presents considerable ARCH effects and the null hypothesis of homoskedasticity must be rejected, at the 5% significance level, for the residuals of OIL also.

As aforementioned, the presence of conditional heteroskedasticity in the cointegration error can cause the DF test and Johansen's LR test to overreject the null hypothesis. However, this problem does not arise in our application of the Engle-Granger cointegration test, since a Phillips-Perron (PP) test, which is robust with respect to heteroskedasticity, was performed on the residuals of the long-run equilibrium equation along with an ADF test. In Section 6.3.1, the results of both tests indicate the presence of a cointegrating vector. Moreover, in Section 6.3.2 the cointegrating relationships and loading matrix were estimated with the R function `ca.jo()` using the ML estimator, which is robust to departures from the standard model assumptions as well as to overparametrization (Gonzalo (1994)). This findings may indicate that, although the variance of some of the single residual series is likely to be inhomogeneous, conditional heteroskedasticity is not significant in the cointegration error of the VECM model, as it is, as a matter of fact, confirmed by the results of the multivariate ARCH-LM test. On these grounds, we hold that the problem of ARCH effects, which mainly concerns the interest rate residual series, appears not to have any significant repercussion on the validity of the standard asymptotic tests used in this application and does hence not nullify the beforehand conducted analysis as well as subsequent operations.

6.4.3 Autocorrelation Test

Residuals	DJX	IP	M1	TB3M	OIL	CPI
Statistic	3.7042	5.3611	3.0679	5.1564	1.6245	6.5672
p-value	0.9882	0.9448	0.995	0.9525	0.9998	0.8848
Squared Residuals	**DJX**	**IP**	**M1**	**TB3M**	**OIL**	**CPI**
Statistic	17.749	4.5823	21.638	85.584	18.207	24.17
p-value	0.1235	0.9705	0.0418	3.51×10^{-13}	0.1096	0.0193
Multivariate	**Portmanteau**			**Adj. Portmanteau**		
Statistic	110.86			114.02		
p-value	2.2×10^{-16}			2.2×10^{-16}		

Table 6.24: Univariate and multivariate serial correlation test

The absence of serial correlation in the residuals of a VAR or VECM model can be tested by means of the multivariate Portmanteau test, described in Section 2.3.3. The test considers the null hypothesis of zero residual autocorrelations at lags 1 to m against the alternative hypothesis that at least one autocorrelation coefficient is non-zero. Besides the classical multivariate Portmanteau statistic from equation (2.10), an adjusted statistic with superior small sample properties is proposed in equation (2.11). As in the univariate Portmanteau test, the number of lags m included in the multivariate test statistic can affect the power of the test. The R function `serial.test()` from the packages `vars` implements the multivariate Portmanteau test. The default is to compute the standard Portmanteau statistic (2.10), but the adjusted version can be chosen by setting the argument `type` equal to `"PT.adjusted"`. The argument `lags.pt=12` specifies that the number of lags that should enter the test statistic is set equal to the frequency of the data. Additionally, the single residual series are tested for zero autocorrelation in lags 1 to 12 by means of the univariate Ljung-Box statistic from Section 2.3.3, which is computed via the R function `Box.test` contained in the package `stats`. The results of the univariate and multivariate autocorrelation test are reported in Table 6.24. As implied by the autocorrelograms in Appendix A, serial correlation appears not to be an issue for any of the residual series. However, when the squared residuals are considered, M1, TB3M and CPI have significant residual autocorrelation at the 5% significance level. This confirms the result of the ARCH-LM test that the residuals of the interest rate series are highly heteroskedastic. The low p-values of the multivariate Portmanteau and adjusted Portmanteau test indicate that serial correlation in the residuals of the level-VAR is significant, which might be a consequence of the presence of ARCH effects in some of the considered time series. However, the results in Table 6.24 clearly indicate that residual autocorrelation is not significant for all of the considered time series. Therefore, the contradictory result of the multivariate test can be explained by the presence of some degree of dependency in the second order moments of the residual distribution, caused by inhomogeneous variance. Since in this case the presence of conditional heteroskedasticity has been shown not to be of such entity as to affect previous analysis (Section 6.4.2), the estimated VECM model is considered adequate for describing the true DGP.

6.5 Structural Analysis and Forecasting

6.5.1 Granger Causality

Formally, the existence of Granger causality from a variable y_{2t} to another variable y_{1t} entails that, for some forecast horizon h, a better forecast of y_{1t} is achieved by including information about past values of y_{2t} to the set of information about y_{1t}. In this section, we intend to test if the selected macroeconomic factors are causal to the U.S. stock market index, such that past values of the macroeconomic variables help improve the forecast of DJX. Besides a Granger causality test, an instantaneous causality test is used to inspect whether adding information about future economic activity to current and past values of the macroeconomic factors has a positive impact on the forecast of DJX. The concept of Granger causality and instantaneous causality in the context of VAR models has been already introduced in Section 4.3. Under the null hypothesis of no Granger causality between the sub-vectors y_{1t} and y_{2t} from a bivariate VAR such as (4.10), $\alpha_{1j} = 0$ for $j = 1, ..., p$, while under the alternative hypothesis, $\exists\, \alpha_{1j} \neq 0,\ j = 1, ..., p$. The absence of instantaneous causality between y_{1t} and y_{2t} is determined by testing that the correlation between the error terms u_{1t} and u_{2t} in (4.10) is different from zero (see Section 4.3). For this purpose, a Wald-type test is used. As it is demonstrated in Lütkepohl (2006), an equivalent representation of instantaneous causality is achieved by placing zero restrictions on $\sigma = \mathrm{vech}(\boldsymbol{\Sigma}_u)$, where vech is the column-stacking operator for symmetrical matrices and $\boldsymbol{\Sigma}_u$ is the variance-covariance matrix of the error process from (4.10). The Wald test for no instantaneous causality tests the null hypothesis $H_0 : \boldsymbol{C}\sigma = 0$ against the alternative $H_1 : \boldsymbol{C}\sigma \neq 0$, were $\boldsymbol{C}$ is a matrix of dimensions $(K \times n(n+1)/2)$ and rank K, which singles out the relevant covariances of u_{1t} and u_{2t}. The Wald statistic is defined as

$$\lambda_w = T\tilde{\sigma}'\boldsymbol{C}'[2\boldsymbol{C}\boldsymbol{D}_n^{+}(\tilde{\boldsymbol{\Sigma}}_u \otimes \tilde{\boldsymbol{\Sigma}}_u)\boldsymbol{D}_n^{+'}\boldsymbol{C}']^{-1}\boldsymbol{C}\tilde{\sigma}$$

where $\boldsymbol{D}_n^{+}$ is the Moore-Penrose inverse of the $(n^2 \times \frac{1}{2}n(n+1))$ duplication matrix $\boldsymbol{D}_n$, $\tilde{\boldsymbol{\Sigma}} = T^{-1}\sum_{t=1}^{T}\hat{u}_t\hat{u}_t'$, n is the number of variables in a VAR(p) and T is the number of observations. The test statistic λ_w is asymptotically distributed as a χ^2 with K degrees of freedom.

The R function `casuality()` from the package `vars` implements a causality test as well as an instantaneous causality Wald-type test. The first argument is an object of class `varest`, which is generated by the function `VAR()`. Hence, a VAR model with 12 lags and an intercept is

	Statistic		p-value	
Variable	Granger	Instant	Granger	Instant
IP–DJX	5.086	0.70133	0.0004889	0.4023
M1–DJX	2.5076	3.3823	0.01515	0.0659
TB3M–DJX	1.8246	0.02080	0.06103	0.8853
OIL–DJX	0.056759	0.25237	0.9448	0.6154
CPI–DJX	0.84984	0.52495	0.467	0.4687
$\boldsymbol{v}_1$–DJX	2.6403	2.2014	0.0001033	0.8206

Note: $\boldsymbol{v}_1 = (\text{IP, M1, TB3M, OIL, CPI})$

Table 6.25: Granger causality from macroeconomic variables to DJX

estimated for the analyzed variables before conducting the Granger causality test, as it is shown in the R code snippet below. The character vector `cause` is used to select the variable string that is believed to be causal to the other variables in the system.

```
dataset <- cbind(DJX, IP)
var.model <- VAR(dataset, p=12, type="constant")
causality(var.model, cause="IP")
```

The results of the test in Table 6.25 show that there is Granger causality from industrial production and money supply to the Dow Jones stock index at the 5% significance level, TB3M Granger causes DJX at the 10% significance level, whereas the contribution of past values of OIL and CPI to the optimal forecast of DJX is not significant.

When the set of macroeconomic variables $\boldsymbol{v}_1 = (IP, M1, TB3M, OIL, CPI)$ is considered, the output of the Granger causality test indicates that past macroeconomic information helps determining future stock prices. On the other hand, the null hypothesis of no instantaneous causality could not be rejected at the 5% significance level, which entails that the optimal forecast of DJX is not sensitive to the inclusion of future information about the set of macroeconomic factors. The results of examining reverse causality from the U.S. stock index to macroeconomic variables, displayed in Table 6.26, suggest that the lagged levels of DJX have some influence

	Statistic		p-value	
Variable	Granger	Instant	Granger	Instant
DJX–IP	13.802	0.7013	8.73×10^{-11}	0.4023
DJX–M1	1.2979	3.3823	0.2488	0.0659
DJX–TB3M	2.3248	0.0208	0.0142	0.8853
DJX–OIL	2.0694	0.2524	0.1271	0.6154
DJX–CPI	2.4512	0.5250	0.0625	0.4687

Table 6.26: Granger causality from DJX to macroeconomic variables

in predicting future values of IP and TB3M, at the 5% significance level. The optimal forecasts of the other variables, however, are not responsive to historical stock prices and instantaneous causality is not detected for any of the considered macroeconomic variables.

6.5.2 Impulse Response Analysis

Impulse response functions are used to quantify the response of one or more variables to an impulse of another variable, caused by an exogenous shock, when both variables are in the same system. As was outlined in Section 4.3.2, when the shocks do not occur in isolation, orthogonal innovations, which are by definition uncorrelated, are used in an impulse response analysis. Since in this application the level-VAR representation of a VECM is considered, the stability condition for VAR(p) processes, outlined in Section 4.3, is not satisfied. In the instance of a nonstationary cointegrated process, shocks may have a permanent effect on the system. Although the Wold moving average representation from (4.24) does not hold for nonstationary VAR models (Lütkepohl and Krätzig (2004)), the impulse response matrices $\mathbf{\Phi}_k$ can still be computed according to (4.22). In this application, the short-term behavior of DJX in response to a shock in the system is of interest. Since we are using orthogonal impulse response functions, the variables composing the system must be ordered from the most to the least exogenous in order for the impulse response function to be meaningful. It is common practice to rely on either economic theory, statistical tests or a combination of both to determine a reasonable ordering

(Lütkepohl (2006)). According to Darby (1982), an increase in the price of crude oil is considered to be one of the main causes of economic recession, due to the fact that oil is an essential input for energy production in almost all sectors. Cologni and Manera (2008) examined the dynamic relationship between the UK Brent price of oil, short term interest rates, money supply, inflation rate and real gross domestic product by estimating a structural VECM and found that the world price of oil significantly influenced the long-term equilibrium relationship between the economic variables. On these grounds, it seems reasonable to consider oil price as the most exogenous variable and DJX as the least exogenous, since the stock price is believed to be directly influenced by the other variables in the system (see Section 6.1 for a discussion). The degree of exogeneity of the remaining variables is determined by means of a Granger causality test. The concept of Granger causality and instantaneous causality is treated in Section 4.3, while the procedure for performing a causality test in R is shown in Section 6.5.1. By testing the null hypothesis that the single time series are not causal to the other variables in the system, we attempt to achieve a reasonable ordering. According to the test results displayed in Table 6.27, the analyzed variables should be ordered according to $\boldsymbol{Y}_t = (\mathrm{OIL}, \mathrm{IP}, \mathrm{TB3M}, \mathrm{CPI}, \mathrm{M1}, \mathrm{DJX})$. Due to the difficulty of establishing the precise placement of TB3M and CPI when the variables are regarded as singularly causal to the others, a causality test has been run on the vectors $\boldsymbol{v}_1 = (\mathrm{OIL}, \mathrm{IP}, \mathrm{TB3M})$ and $\boldsymbol{v}_2 = (\mathrm{OIL}, \mathrm{IP}, \mathrm{CPI})$, whose result is shown in Table 6.27.

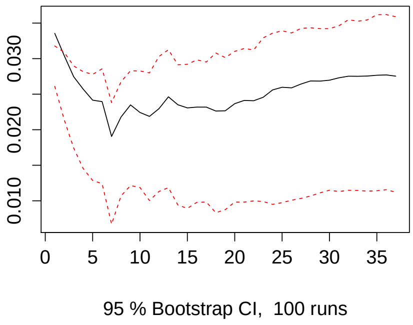
An impulse response analysis is implemented in R with the function `irf()` from the package `vars`. The character vectors `impulse` and `response` specify which variables should be considered as the impulses and responses, respectively. By default, the argument `ortho=TRUE` computes orthogonal impulse responses, whose output is influenced by the ordering of the variables in the time series vector $\boldsymbol{Y}_t$. The significance of the impulse response analysis can be assessed by estimating a confidence interval for the IRF via the function `boot=TRUE`, which by default returns a 95% confidence band, plotted as a dotted red line in Figure 6.3. The length of the impulse response, which is controlled by the argument `n.ahead`, is set to 3 years. Figure 6.3 shows the responses of DJX to a unit shock in each of the variables. The estimated impulse response functions and the relative confidence bands are stored in Tables B.3–B.8 in Appendix B. In Figure (a), the DJX series is seen to respond to a one standard deviation shock in its own value with a very sharp decrease from month 1 to 7, where it reaches a minimum of 0.019 units. A succession of higher and lower values can be noticed during months 7 to 14, after which the

	Statistic		p-value	
Variable	Granger	Instant	Granger	Instant
IP	1.9821	14.128	0.0000172	0.01481
M1	0.85991	12.346	0.7683	0.03034
TB3M	1.67	9.4709	0.001201	0.09169
OIL	2.7564	65.673	5.65×10^{-11}	8.12×10^{-13}
CPI	1.4766	70.782	0.01141	7.05×10^{-14}
(OIL, IP, TB3M)	2.953	76.101	2.2×10^{-16}	9.56×10^{-13}
(OIL, IP, CPI)	1.7815	20.954	0.0000037	0.01286

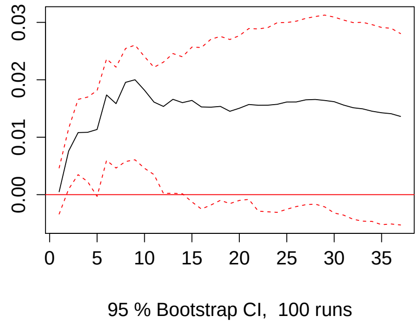
Table 6.27: Granger causality test

series displays a more regular increasing pattern until it reaches a value of 0.027 in year 3 — 0.007 points below its initial value. A shock in IP (Figure (b)) is immediately followed by an increase in the value of DJX, which peaks after 9 months, reaching a maximum of 0.020, drops to a value of 0.015 in month 19 and then gradually starts rising again before settling for a value of 0.014 in year 3. A one-unit innovation in M1 (Figure (c)) prompts a 0.002 units drop in the value of DJX during the first two months, which is followed by a steep increase that peaks in month 9, where a value of 0.013 is reached. Starting from period 25, the value of DJX begins to rise in a steady fashion and at a much slower rate, until it becomes 0.015 in year 3.

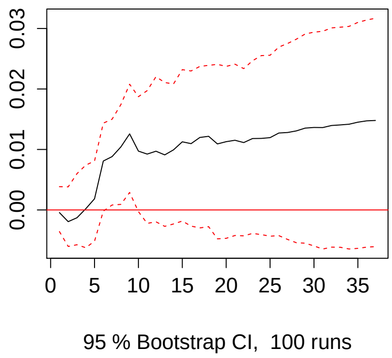
A decline in DJX, whose value became negative for the first 10 months, follows an impulse in the short term interest rates, as shown in Figure (d). The decrease eventually reversed in month 11, the series grew to a maximum of 0.014 after 20 months, and began descending again towards zero afterwards. After an initial 0.005 units drop in period 2, followed by a small increase in periods 3 and 4, the value of DJX started to decline at a reasonably fast pace in response to a shock in the price of oil, as displayed in Figure (e). After reaching a minimum of -0.023 in periods 16 and 17, the value of the stock index diverted its path and began moving upwards until it reached a value of -0.022 in the last month of the IRF. A one-unit shock in the rate of inflation clearly had a negative effect on the stock price index, driving its value below



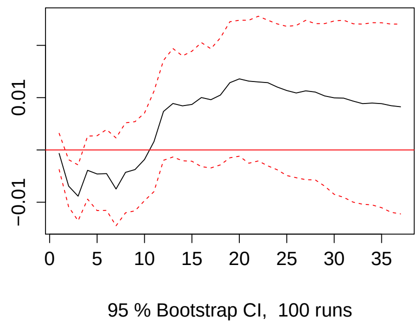
(a) Impulse of DJX



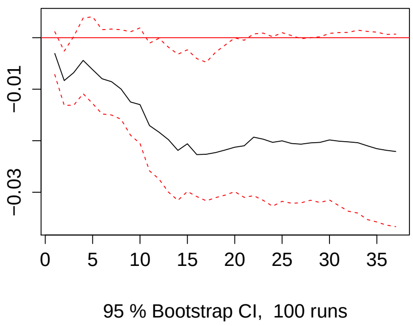
(b) Impulse of IP



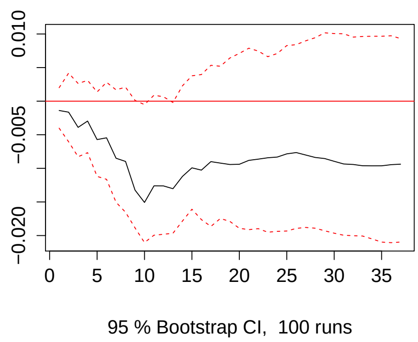
(c) Impulse of M1



(d) Impulse of TB3M



(e) Impulse of OIL



(f) Impulse of CPI

Figure 6.3: Impulse responses of DJX

zero immediately after the first period (Figure (f)). A rather sharp decrease can be noticed from month 1 to 10, followed by an alternation of higher and lower values between periods 11 and 19, after which the DJX series started to decrease at an approximately constant rate until the end of the third year, where its value added up to -0.009. In accordance with the nonstationary nature of the analyzed time series, shocks appear to have a permanent effect on the system and the response variable is not seen to revert back to the original value after an innovation occurred. Roughly, the confidence bands follow the same direction as the impulse response functions for all series, indicating that the IRF is quite stable, although there are occasional widenings of the confidence intervals around the IRF. For the most, the confidence interval is relatively tight around the IRF during the first year of the analysis, and starts becoming more outspread afterward. Periods in which the confidence bands do not follow the IRF closely suggest a lack of accuracy in the impulse response analysis as well as a loss in predictive power. This is confirmed by the estimated values of the confidence intervals in Appendix B.2, which are particularly outstretched in year 2 and 3.

6.5.3 Forecast Error Variance Decomposition

The forecast error variance decomposition (FEVD), described in Section 4.3.3, is used to assess the contribution of innovations in variable j to the h-step forecast error variance of variable i. The calculated value of the FEVD is a percentage figure, since the FEVD results from the quotient between the elements of the squared orthogonal impulse response matrix $\mathbf{\Psi}_k$ (Section 4.3.2) and the forecast error variance of $y_{i,T+h}$, as in equation (4.26). The R function `fevd()` from the package `vars` implements the FEVD analysis. In this example, the argument `n.ahead` is set equal to 36, so that the FEVD is computed for a 3 years ahead forecast. The graphical results are displayed in Figures 6.4 and 6.5, whereas the calculated values of the FEVD are shown in Tables B.9–B.14 in Appendix B.

The forecast errors of the considered time series are mostly attributable to own innovations, with the only exception of CPI. As it is shown in Figure 6.4, the largest contribution to the forecast error variance of the Dow Jones index stems from shocks in industrial production and the price of oil, with a maximum of 18% and 13%, respectively. On average, IP is responsible for 15% of the variance in the forecast error of DJX, while OIL accounts for 9% of it. The contribution

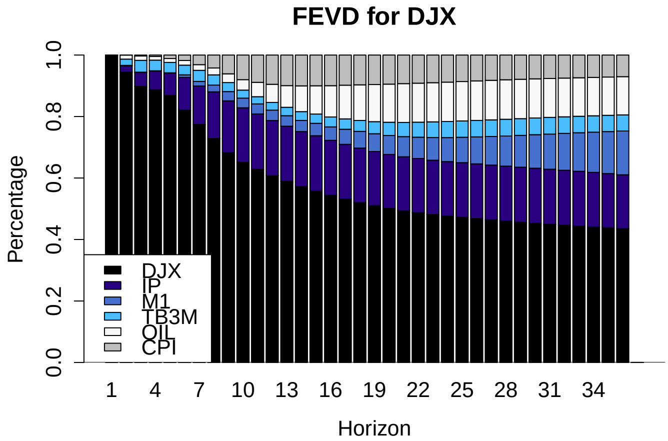


Figure 6.4: Forecast error variance decomposition of DJX

of money supply did not exceed 6% during the first half of the observation period, but increased dramatically during the second part of the forecast horizon and reached a value of 14% in year 3. On average, 7% of U.S. stock price variability is caused by CPI, 6% is caused by M1 and 4% is caused by TB3M. Hence, over the 3 years of the FEVD, shocks in macroeconomic variables were accountable for an average of about 40% of the variance in DJX. The stock index plays a large role in determining the forecast error variance of industrial production: up to a maximum of 16% of the variance in the forecast of IP is due to shocks in DJX, while at most 14% is due to shocks in M1. On average, DJX contributes to explaining 12% of the variance in IP, while the contribution of TB3M, OIL and CPI never exceeds 5%. On average, 14% of the variance in the forecast error of money supply is explained by shocks in industrial production, about 20% is jointly explained by shocks in TB3M, OIL and CPI, whereas DJX accounts for less than 1% of the variability in M1, as is it shown in Figure 6.5. Own innovations appear to be the main explanatory factor of the forecast error in TB3M during the entire horizon; the percentage of variance which is due to TB3M never goes below 83% and is 86% on average. The other main explanatory factor of the variance $\sigma^2_{\mathrm{TB3M}}(3)$ is industrial production with a mean value of 9%, while less than 5% of the interest rate volatility is due to shocks in DJX, CPI, M1 and OIL [8]. In Figure 6.5, the oil price seems to be responsible for the largest part of its forecast error variance,

[8] The notation $\sigma^2_{\mathrm{TB3M}}(3)$ indicates the 3-step ahead forecast error variance of TB3M and is used in Section 4.3.3.

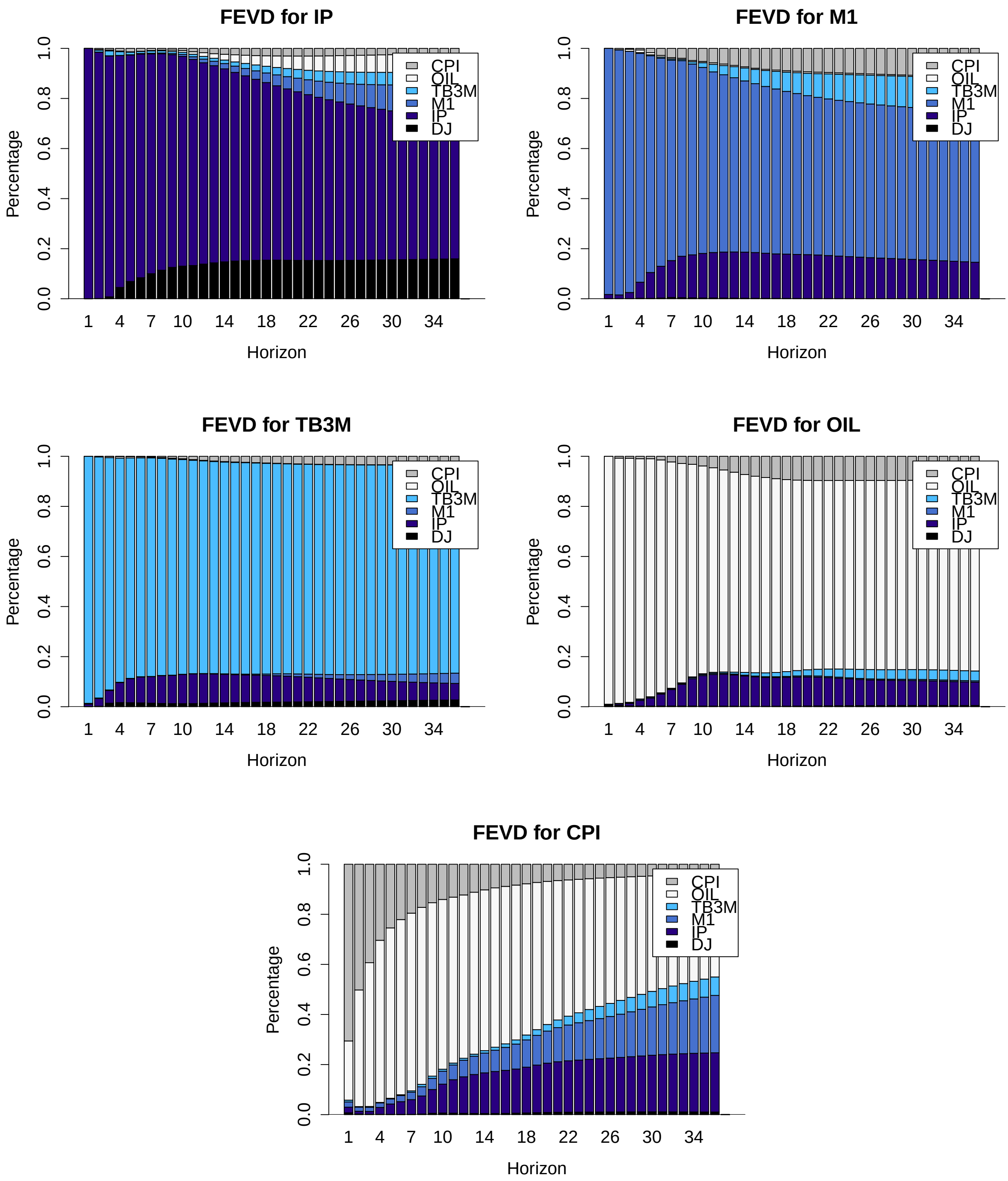


Figure 6.5: Forecast error variance decomposition of macroeconomic factors

with an average value of 81% and a minimum value of 75%. The other main contributors are IP, CPI and TB3M, with 9%, 6% and 2%, on average. The Dow Jones stock index and money supply are not significantly involved in the determination of the variance $\sigma^2_{\mathrm{OIL}}(3)$ and their contribution is lower than 1%. Of all variables, the consumer price index is the least involved in explaining its own forecast error variance. Its contribution is only 13% on average and reaches a minimum of 4% in year 3. The proportion of the variance in the forecast error of CPI that is caused by shocks in the OIL price adds up to 56% on average and has a maximum value of 71%. About 15% of the variance in CPI is explained by IP, 11% is explained by M1, and less than 1% is explained by DJX. On the grounds of this analysis, it can be concluded that the result of the FEVD for DJX support the argument that the analyzed macroeconomic variables help explaining stock market movements to a certain extent.

6.5.4 Forecasting

The existence of Granger causality from the selected macroeconomic variables to the U.S. stock market index, established in Section 6.5.1, suggests that the estimated VECM may be an efficient tool for forecasting purposes. The calculation of forecasts and forecast confidence intervals for a VAR model with either known or estimated parameters was treated in Section 4.3.1. In order to assess the forecasting ability of the VECM from Section 6.3.2, an out-of sample forecast is performed for the analyzed time series via the R function `predict()`, included in the package `vars`. In particular, a 3-years ahead forecast is conducted from 2012 to 2015, based on the data from the sample from 1989 to 2012. The length of the forecast horizon is selected via the argument `n.ahead` and a 95% confidence interval for the forecasted series is computed by default. Note that the function `predict()` does not provide for a correction of the estimation error in the VAR model coefficients, resulting in smaller confidence intervals than otherwise.

In Figures 6.6 and A.9 (Appendix A), the forecasts of the analyzed time series are compared to their observed time path. The values of the forecasts, along with the observed values of each time series and the confidence intervals, are stored in Appendix B.4. In general, a VAR model is considered adequate for forecasting purposes when all its observed values are within the forecast interval (Lütkepohl (2006)). In this application, the observed values of DJX, IP, M1 and TB3M lie within the 95% forecast interval for the whole forecast horizon, while the

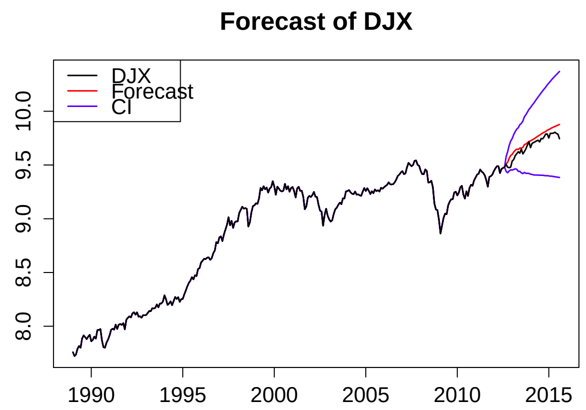


Figure 6.6: Out-of-sample forecast of DJX

realizations of OIL and CPI trespass the confidence bands 27 and 14 months after the forecast origin, respectively (Figure A.9, Appendix A). As it is shown in Figure 6.6, the VECM provides a reasonably accurate forecast of the Dow Jones price series, since the discrepancies between the forecast and the actual realizations of the stock price series are of small size. This leads us to conclude that the selected macroeconomic factors possess a remarkable explicative power over the stock market and are helpful predictors of future stock price movements when they are considered in the framework of a cointegrated VECM.

Conclusions

This study was aimed at examining the long-run relationship between the U.S. stock market, represented by the Dow Jones Industrial Average, and a set of selected macroeconomic variables composed of industrial production, money supply, short-term interest rates, crude oil price and the consumer price index as a proxy for inflation. To this purpose, the methodological framework of cointegration analysis and a vector error correction model (VECM) were employed. The data-set used in the analysis consists of monthly observations of the stock index and the macroeconomic factors collected during the period from January 1989 to August 2015. The results of the cointegration test and the vector error correction model indicate that the stock price index is cointegrated with the set of macroeconomic variables, implying that a long-run equilibrium relationship between the U.S. stock market index and the selected macroeconomic factors exists such that information about relevant economic indicators is reflected in stock prices. The VECM demonstrates the existence of a simultaneous and interactive relation between the economic variables and the stock price index, showing that the U.S. stock market signals changes in economic activities and these changes are significantly priced in the stock market index. The effort of determining the existence of a long-term linkage between the Dow Jones index and the single macroeconomic variables by means of an Engle-Granger test unfolded the existence of a significant relationship between the stock index and industrial production only. Nevertheless, Johansen's cointegration test applied to the system of variables revealed that the U.S. stock index forms a cointegrating relationship with the selected economic variables and the presence of one cointegrating vector in the data-set was detected by Johansen's maximum eigenvalue test. In an attempt to assess the contribution of the single variables to the cointegrating vector, a likelihood ratio (LR) test of linear restrictions was performed, whose output revealed a non-significant coefficient for the price of oil and inflation and consequently, a lack of participation to the error correction mechanism. The significant long-run coefficients of industrial production, money supply and interest rates, however, illustrate that these variables respond to short-term deviations from the equilibrium relationship, which is confirmed by the statistically significant speed of adjustment coefficient associated with the first eigenvector in the VECM equation for the Dow Jones index. Additionally, the hypothesis that past changes in macroeconomic variables do not influence the current value of the stock index, i.e. that the coefficients of lagged differences and the error correction term are zero, is rejected in our empirical analysis. The estimate of the VECM illustrates that the error correction term and some of the lagged differences are significant in the Dow Jones equation, showing that there are meaningful short-run dynamic effects such that the stock index adjusts to the previous equilibrium error. Consistently with economic theory and previous empirical surveys, it is observed that the stock price is positively influenced by industrial production, 3-month Treasury bill rates, oil price and money supply, in the long run, while it is negatively related to inflation. The adequacy of the estimated VECM for describing the DGP of the underlying data-set was verified by testing the model residuals for normality, heteroskedasticity and serial correlation. Although the hypothesis of non-normality could not be rejected, the multivariate ARCH-LM test and the Ljung-Box test for residual autocorrelation did not highlight any significant violations of the standard model assumptions, confirming the validity of the VECM for analysis purposes. The results of examining whether the single macroeconomic variables Granger cause the U.S. stock index showed that the optimal forecast of the Dow Jones index is sensitive to the inclusion of historical values of industrial production, money supply and interest rates, while it is not influenced by past values of oil price and inflation as well as by information about future economic activity, establishing the absence of instantaneous causality. It has been shown that the selected macroeconomic factors are significant indicators of stock price movements; however, it appears from inspecting the reverse causality from the stock market to the macroeconomic variables that the Dow Jones index also has some explicative power over industrial production, interest rates and inflation, while future stock price levels are shown to influence the optimal forecast of money supply. The existence of significant short-run dynamic effects between the U.S. stock index and the set of macroeconomic variables was inspected by performing an impulse response analysis, whose output suggests a positive short-term response of the Dow Jones index to an exogenous shock in industrial production and money supply, a negative response to innovations in inflation and the price of oil and mixed reactions to a shock in interest rates. In accordance with the nonstationary nature of the analyzed data-set, the effect of an exogenous shock on the system is not seen to die out quickly and the stock market index does not revert back to the original equilibrium value after an innovation occurred, at least in the short run. A forecast error variance decomposition was realized in order to examine the contribution of shocks in macroeconomic activity to the forecast error variance of the stock index. The results of the analysis revealed that, over a period of 3 years, 15% of the variance in the forecast error of the Dow Jones index is caused by innovations in industrial production, 9% is caused by innovations in the price of oil, 7% is caused by inflation, 6% is caused by money supply and 4% can be attributed to innovations in interest rates, on average. The fact that about 40% of the forecast error variance of the stock index can be explained by shocks in macroeconomic variables indicates that U.S. stock price variability is fundamentally linked to changes in economic variables. In the context of an out-of-sample 3 years ahead forecast based on the sample period 1989–2012, the forecasting ability of the estimated VECM was tested. The adequacy of the model for forecasting purposes is established by the accuracy of the forecasted U.S. stock price series and the hypothesis that the selected macroeconomic variables possess a certain explanatory power over stock market movements and are relevant factors in the determination of future stock prices is validated.

Appendix A

Graphical Results

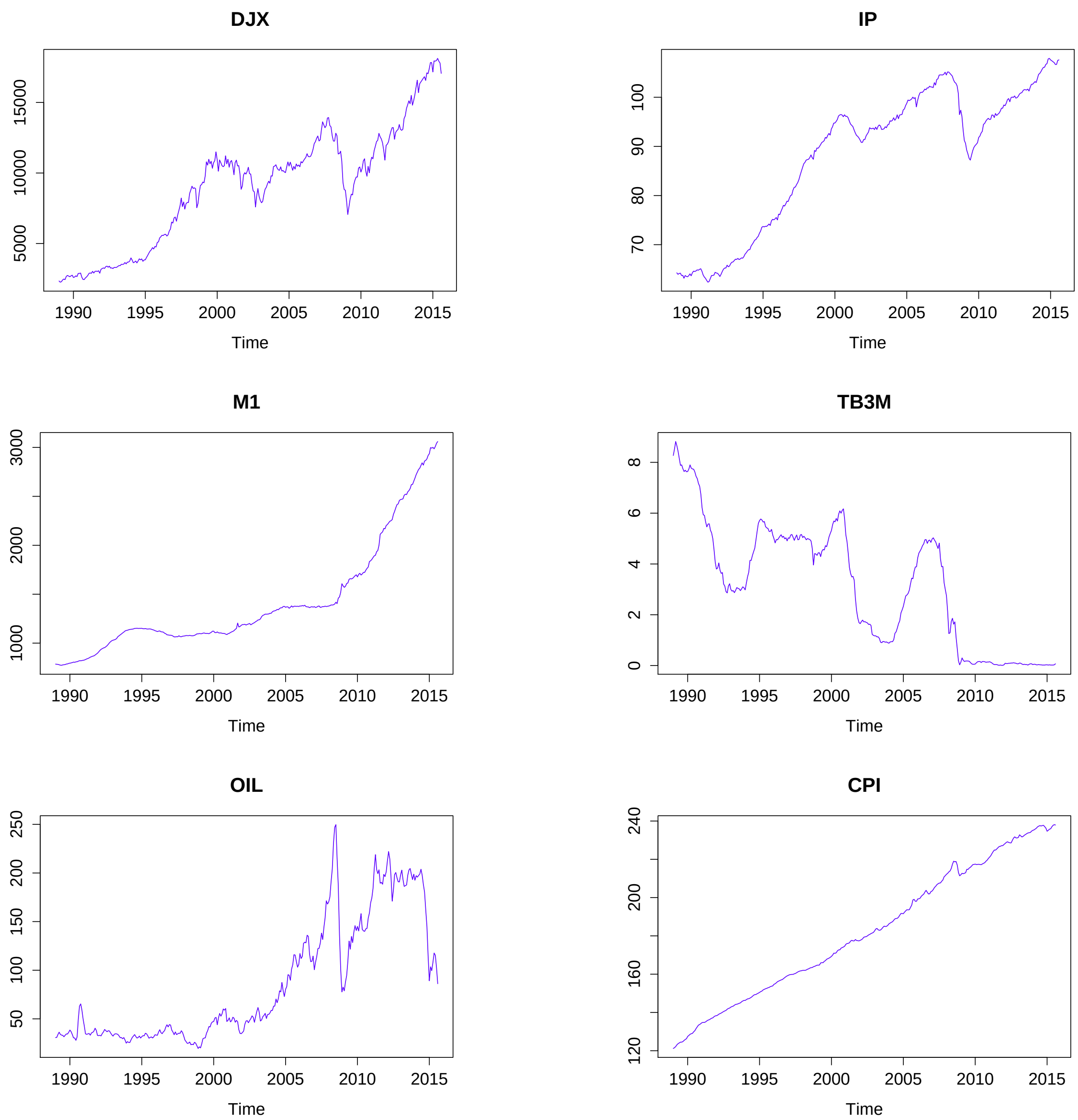
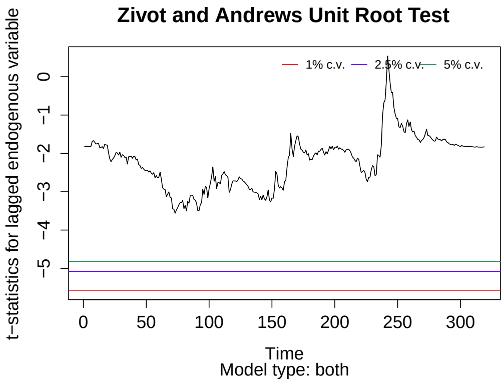
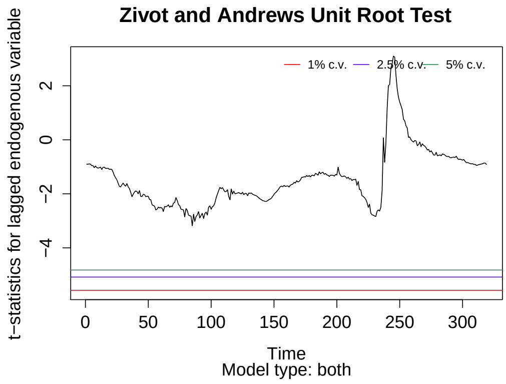


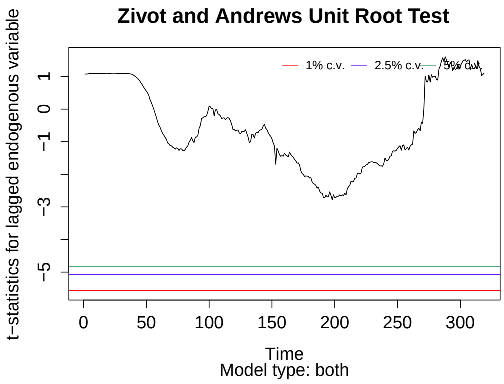
Figure A.1: Time plot of the untransformed series



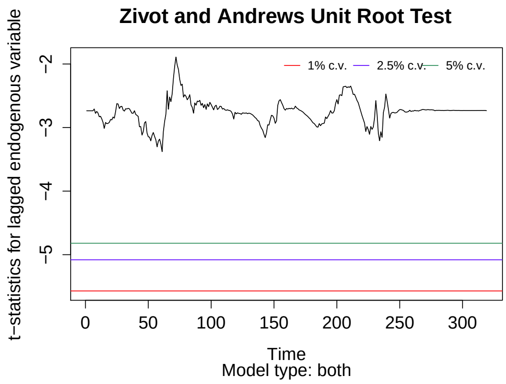
(a) DJX



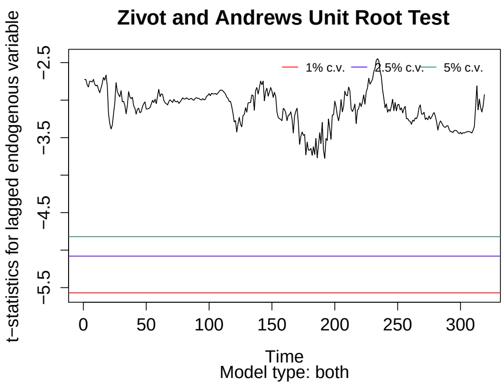
(b) IP



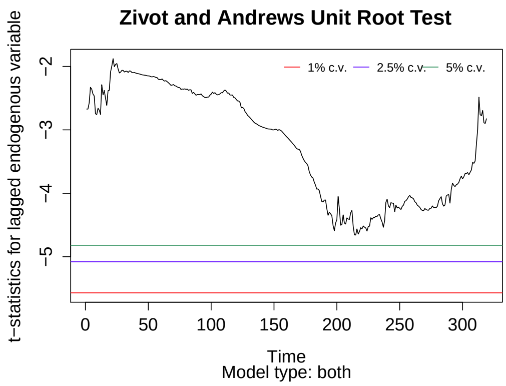
(c) M1



(d) TB3M



(e) OIL



(f) CPI

Figure A.2: Andrews-Zivot test

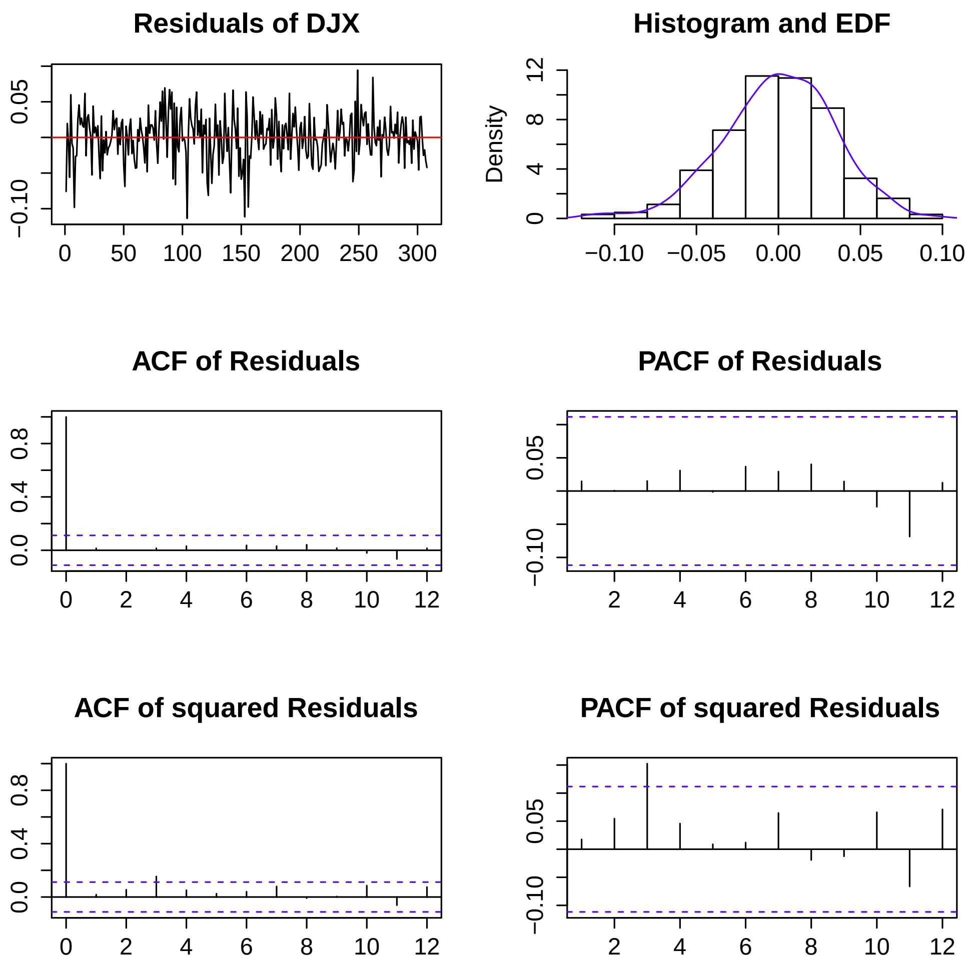


Figure A.3: Residual diagnostic of DJX

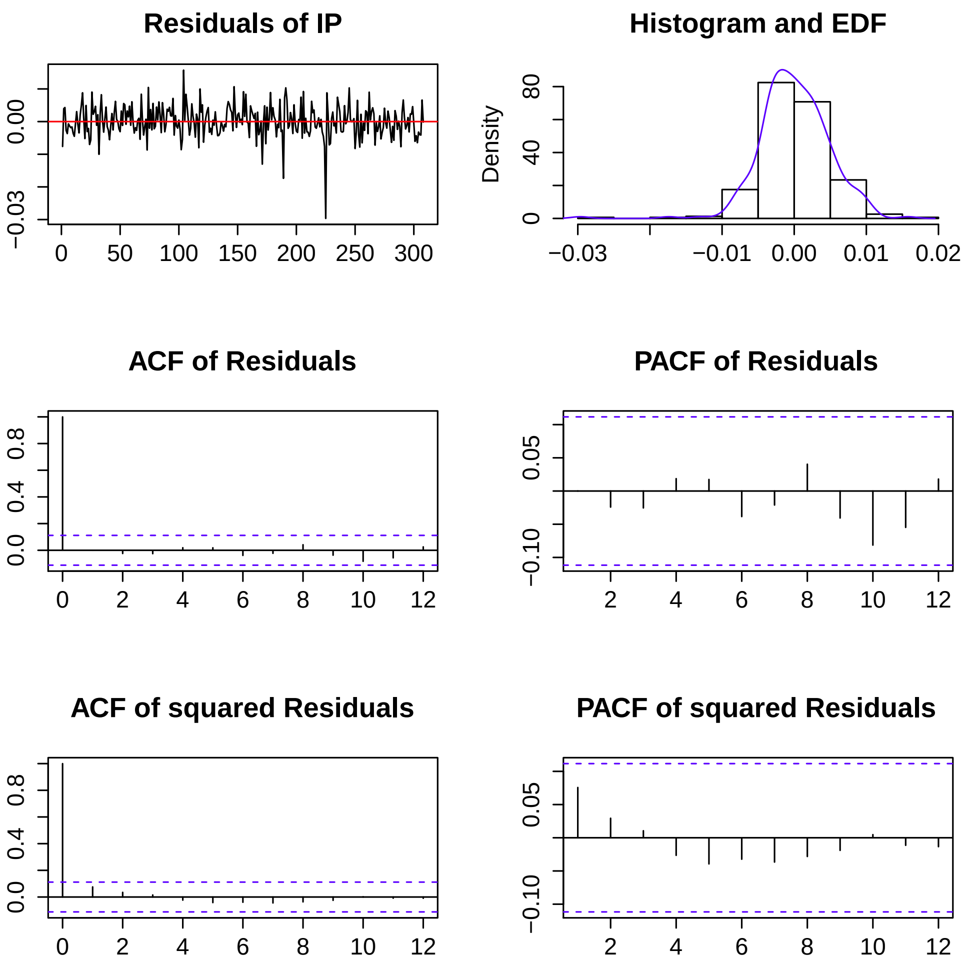


Figure A.4: Residual diagnostic of IP

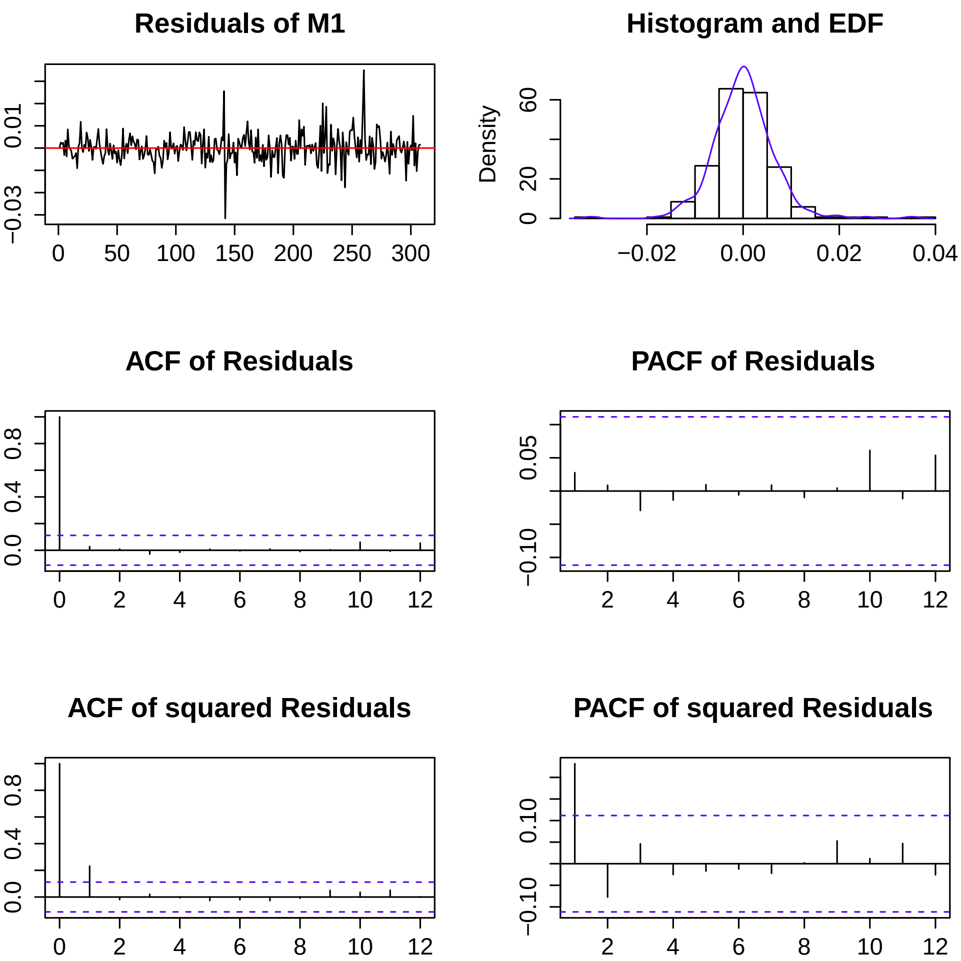


Figure A.5: Residual diagnostic of M1

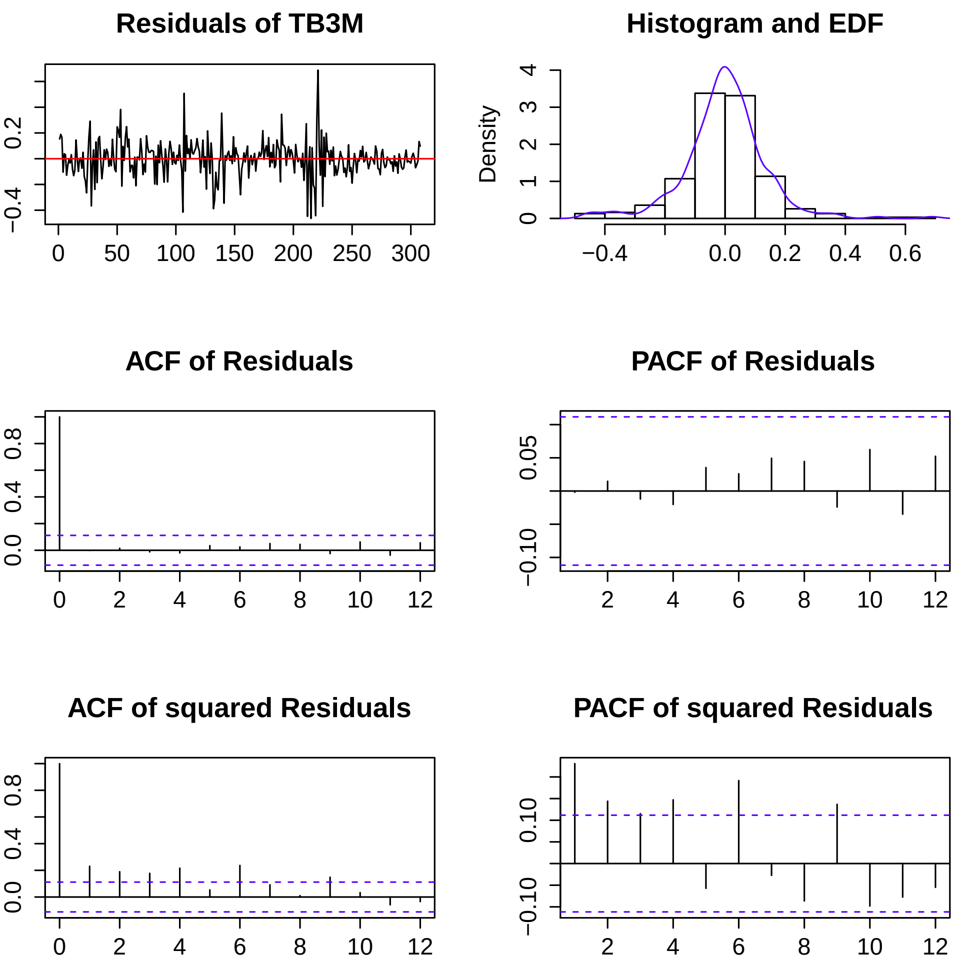


Figure A.6: Residual diagnostic of TB3M

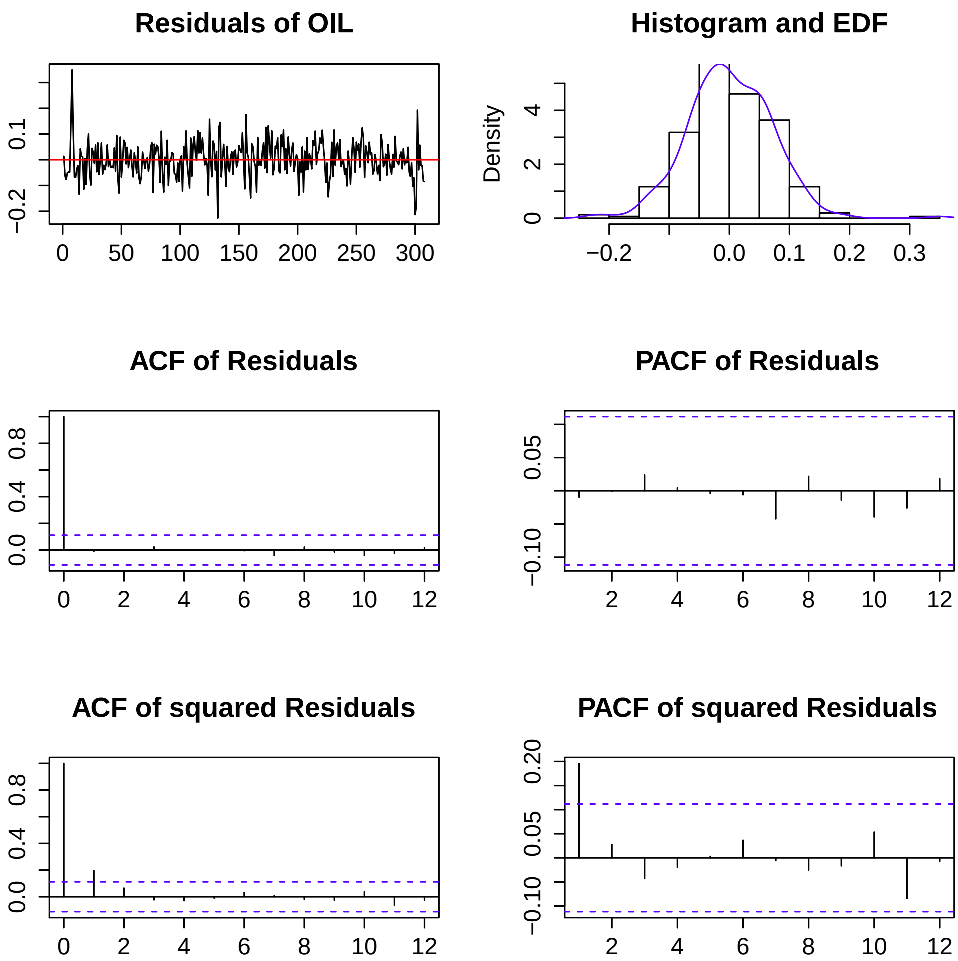


Figure A.7: Residual diagnostic of OIL

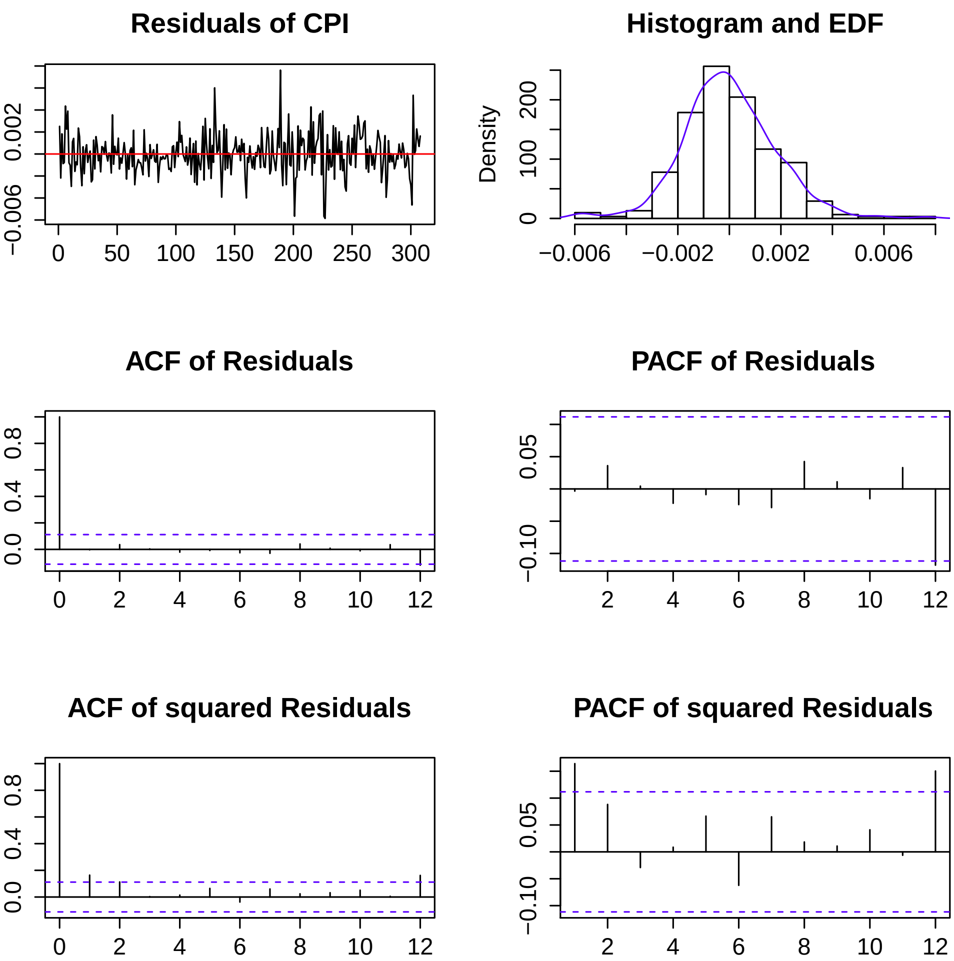


Figure A.8: Residual diagnostic of CPI

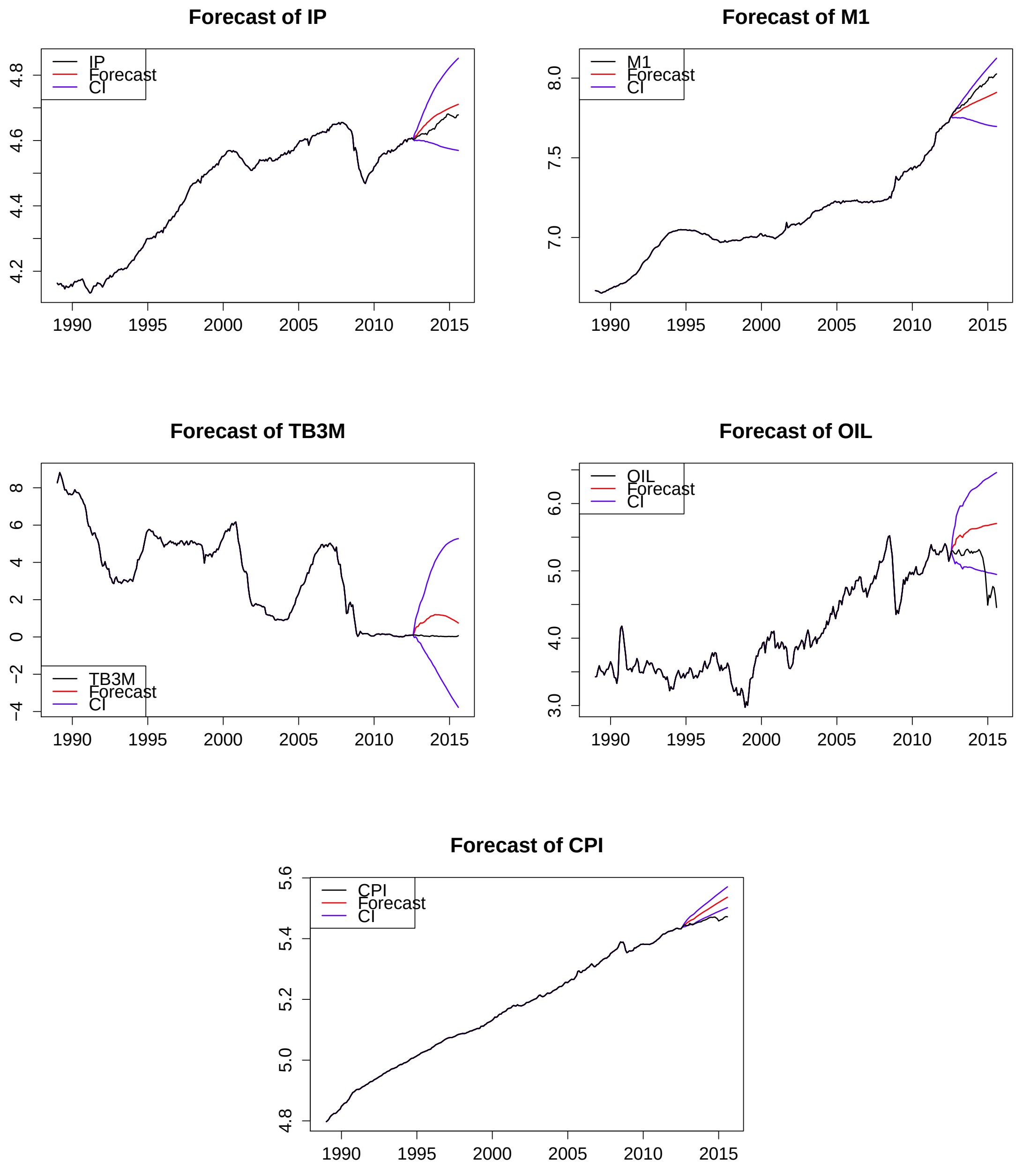


Figure A.9: Out-of-sample forecast of macroeconomic factors

Appendix B

Tables of Results

Vector Error Correction Model

Table B.1: VECM for DJX

Variable	Estimate	Std. Error	t-value	Pr(>|t|)
ΔDJX_{t-1}	0.0319821	0.0659554	0.4849052	0.6281859
ΔIP_{t-1}	1.1465510	0.4807900	2.3847229	0.0178695
$\Delta M1_{t-1}$	−0.3042460	0.3285425	−0.9260474	0.3553516
$\Delta TB3M_{t-1}$	−0.0551654	0.0162485	−3.3951002	0.0008026
ΔOIL_{t-1}	−0.0811189	0.0360335	−2.2512064	0.0252779
ΔCPI_{t-1}	0.0527777	1.4088442	0.0374617	0.9701480
ΔDJX_{t-2}	0.0435315	0.0669810	0.6499090	0.5163726
ΔIP_{t-2}	0.6956373	0.4947078	1.4061579	0.1609706
$\Delta M1_{t-2}$	−0.0197584	0.3278311	−0.0602700	0.9519908
$\Delta TB3M_{t-2}$	−0.0216771	0.0174657	−1.2411256	0.2157711
ΔOIL_{t-2}	0.0639366	0.0399661	1.5997697	0.1109652
ΔCPI_{t-2}	−2.3955259	1.4564257	−1.6447979	0.1013206
ΔDJX_{t-3}	0.0914785	0.0680346	1.3445886	0.1800275
ΔIP_{t-3}	−0.2885297	0.4930937	−0.5851419	0.5590021
$\Delta M1_{t-3}$	0.1284715	0.3227972	0.3979946	0.6909880

ΔTB3M$_{t-3}$	0.0269467	0.0174731	1.5421828	0.1243466
ΔOIL$_{t-3}$	0.0362581	0.0391849	0.9253084	0.3557350
ΔCPI$_{t-3}$	0.8944347	1.4680298	0.6092756	0.5429181
ΔDJX$_{t-4}$	-0.0271752	0.0700561	-0.3879064	0.6984291
ΔIP$_{t-4}$	-0.4929328	0.5045803	-0.9769163	0.3295943
ΔM1$_{t-4}$	0.3550148	0.3278252	1.0829393	0.2799222
ΔTB3M$_{t-4}$	-0.0160806	0.0181843	-0.8843122	0.3774127
ΔOIL$_{t-4}$	0.0066085	0.0400328	0.1650766	0.8690227
ΔCPI$_{t-4}$	-3.1929453	1.4890033	-2.1443508	0.0330096
ΔDJX$_{t-5}$	0.0032384	0.0699191	0.0463164	0.9630966
ΔIP$_{t-5}$	0.7723521	0.5019502	1.5387027	0.1251943
ΔM1$_{t-5}$	1.1263978	0.3273645	3.4408065	0.0006839
ΔTB3M$_{t-5}$	0.0038492	0.0177081	0.2173695	0.8281051
ΔOIL$_{t-5}$	0.0155443	0.0401890	0.3867795	0.6992622
ΔCPI$_{t-5}$	-0.6722318	1.4547198	-0.4621040	0.6444254
ΔDJX$_{t-6}$	-0.1074813	0.0679296	-1.5822467	0.1149101
ΔIP$_{t-6}$	-0.4866436	0.5053903	-0.9629064	0.3365636
ΔM1$_{t-6}$	0.1145940	0.3254810	0.3520760	0.7250903
ΔTB3M$_{t-6}$	-0.0530174	0.0176560	-3.0027960	0.0029580
ΔOIL$_{t-6}$	0.0257269	0.0399648	0.6437389	0.5203594
ΔCPI$_{t-6}$	-2.7489042	1.4673952	-1.8733224	0.0622393
ΔDJX$_{t-7}$	0.1221818	0.0678719	1.8001827	0.0730875
ΔIP$_{t-7}$	0.8316570	0.5068069	1.6409741	0.1021126
ΔM1$_{t-7}$	0.4320244	0.3275496	1.3189590	0.1884399
ΔTB3M$_{t-7}$	0.0135290	0.0181039	0.7472981	0.4556152
ΔOIL$_{t-7}$	0.0591002	0.0405134	1.4587817	0.1459327
ΔCPI$_{t-7}$	-1.3700052	1.4816598	-0.9246422	0.3560808
ΔDJX$_{t-8}$	0.0615284	0.0676185	0.9099348	0.3637698
ΔIP$_{t-8}$	0.3631409	0.5037103	0.7209321	0.4716529
ΔM1$_{t-8}$	0.4583730	0.3246921	1.4117158	0.1593288

ΔTB3M$_{t-8}$	-0.0021290	0.0179365	-0.1186965	0.9056151
ΔOIL$_{t-8}$	0.0858023	0.0408074	2.1026169	0.0365412
ΔCPI$_{t-8}$	-4.7491641	1.4745094	-3.2208437	0.0014549
ΔDJX$_{t-9}$	0.0129937	0.0667753	0.1945887	0.8458795
ΔIP$_{t-9}$	-0.1610417	0.4726544	-0.3407177	0.7336142
ΔM1$_{t-9}$	0.0100405	0.3116104	0.0322214	0.9743223
ΔTB3M$_{t-9}$	0.0222797	0.0175173	1.2718664	0.2046516
ΔOIL$_{t-9}$	0.0871205	0.0409756	2.1261564	0.0345114
ΔCPI$_{t-9}$	-1.7682170	1.4992050	-1.1794364	0.2393923
ΔDJX$_{t-10}$	-0.0284766	0.0647989	-0.4394608	0.6607231
ΔIP$_{t-10}$	-0.9474843	0.4718728	-2.0079233	0.0457723
ΔM1$_{t-10}$	-0.2102656	0.3149620	-0.6675903	0.5050366
ΔTB3M$_{t-10}$	0.0139669	0.0172256	0.8108251	0.4182694
ΔOIL$_{t-10}$	-0.0517784	0.0409461	-1.2645500	0.2072593
ΔCPI$_{t-10}$	-0.9541594	1.4734578	-0.6475648	0.5178854
ΔDJX$_{t-11}$	-0.0039194	0.0640912	-0.0611529	0.9512884
ΔIP$_{t-11}$	-0.2608196	0.4575072	-0.5700886	0.5691509
ΔM1$_{t-11}$	0.2170435	0.3153972	0.6881593	0.4920171
ΔTB3M$_{t-11}$	0.0025740	0.0163016	0.1578974	0.8746704
ΔOIL$_{t-11}$	0.0886585	0.0414910	2.1368144	0.0336247
ΔCPI$_{t-11}$	-2.1883986	1.2671270	-1.7270555	0.0854443

B.1 VAR Model in Levels

Table B.2: Level-VAR representation of VECM

Variable	DJX	IP	M1	TB3M	OIL	CPI
ΔDJX$_{t-1}$	0.9066513	1.5930255	−0.2418769	−0.0471804	−0.0779599	−0.0749080
ΔIP$_{t-1}$	0.0071386	0.9572864	0.0143214	0.0045752	0.0002959	0.3041546
ΔM1$_{t-1}$	0.0047637	−0.0318564	0.9727415	−0.0004026	0.0033055	−0.4554129
ΔTB3M$_{t-1}$	0.4647997	4.3817119	0.1107167	1.3362003	0.1022029	3.6006534
ΔOIL$_{t-1}$	0.0397886	0.8857655	−0.0855667	−0.0152129	1.3827288	−6.4836238
ΔCPI$_{t-1}$	0.0073103	0.0137052	0.0056905	0.0009318	0.0145767	1.1437803
ΔDJX$_{t-2}$	0.0115494	−0.4509138	0.2844876	0.0334883	0.1450555	−2.4483036
ΔIP$_{t-2}$	0.0056936	0.1857665	−0.0168187	−0.0022387	−0.0040008	−0.2898696
ΔM1$_{t-2}$	−0.0086804	−0.0894097	0.0444209	0.0014609	−0.0161772	0.9700084
ΔTB3M$_{t-2}$	0.1614771	−3.0969026	−0.3860334	−0.3717272	−0.0498586	−6.9744896
ΔOIL$_{t-2}$	0.1266821	−0.4989582	−0.2381794	0.0586563	−0.3746889	10.2806822
ΔCPI$_{t-2}$	−0.0053566	0.0530032	0.0132621	0.0004685	−0.0143433	−0.2264857
ΔDJX$_{t-3}$	0.0479470	−0.9841670	0.1482299	0.0486238	−0.0276785	3.2899606
ΔIP$_{t-3}$	0.0274399	0.0484525	−0.0343454	−0.0064640	−0.0050539	0.2357711
ΔM1$_{t-3}$	−0.0027302	−0.1679236	0.1125611	−0.0065865	0.0040162	−0.0640603
ΔTB3M$_{t-3}$	−0.8262902	0.6656406	1.1408976	0.2836868	0.0628688	−1.4892289
ΔOIL$_{t-3}$	−0.0721292	1.3097608	0.0590737	−0.0812850	0.0392022	−7.9877012
ΔCPI$_{t-3}$	−0.0037408	−0.0194995	0.0012423	−0.0041405	0.0055467	−0.1081179
ΔDJX$_{t-4}$	−0.1186537	−0.2044030	0.2265433	−0.0430273	−0.0296496	−4.0873800
ΔIP$_{t-4}$	−0.0248723	−0.1162195	0.0259733	0.0019951	0.0033062	−0.1318094
ΔM1$_{t-4}$	0.0243812	0.1357587	−0.2383032	0.0076535	0.0048974	0.2672539
ΔTB3M$_{t-4}$	−0.1417868	−4.8612402	−0.9705920	−0.2828924	−0.2919757	5.2938578
ΔOIL$_{t-4}$	−0.1881070	−1.9820424	1.3033724	0.0338444	−0.1906826	6.4085300
ΔCPI$_{t-4}$	−0.0027237	−0.0524578	−0.0088132	0.0020282	−0.0064266	0.2987169
ΔDJX$_{t-5}$	0.0304136	1.2652848	0.7713829	0.0199298	0.0089358	2.5207135
ΔIP$_{t-5}$	−0.0050749	−0.0915181	0.0457749	0.0035946	0.0077415	−0.3563153

$\Delta\mathrm{M1}_{t-5}$	-0.0154320	0.1214152	0.1154087	-0.0026760	-0.0024924	-0.3759436
$\Delta\mathrm{TB3M}_{t-5}$	-0.0535758	0.9466862	0.1508376	-0.0511483	0.0304185	-2.3495756
$\Delta\mathrm{OIL}_{t-5}$	0.0167486	1.2406769	-1.0779556	-0.0432567	0.1807473	-8.2879136
$\Delta\mathrm{CPI}_{t-5}$	-0.0004148	0.0011441	-0.0020966	0.0007211	0.0044328	-0.3264725
$\Delta\mathrm{DJX}_{t-6}$	-0.1107197	-1.2589957	-1.0118037	-0.0568666	0.0101826	-2.0766723
$\Delta\mathrm{IP}_{t-6}$	0.0090904	0.0688702	-0.0565136	0.0045449	-0.0028795	0.0438805
$\Delta\mathrm{M1}_{t-6}$	-0.0002013	-0.0577823	0.0827954	0.0028013	-0.0167824	0.0888895
$\Delta\mathrm{TB3M}_{t-6}$	0.3736393	0.8604236	-0.4645220	0.2240416	0.3382246	-0.5252445
$\Delta\mathrm{OIL}_{t-6}$	0.0376949	-0.7117733	0.8311309	0.0570476	-0.0032989	3.6225113
$\Delta\mathrm{CPI}_{t-6}$	-0.0005904	-0.0186034	-0.0061893	-0.0002852	0.0000284	0.1619895
$\Delta\mathrm{DJX}_{t-7}$	0.2296631	1.3183006	0.3174304	0.0665464	0.0333732	1.3788990
$\Delta\mathrm{IP}_{t-7}$	-0.0103916	-0.1809603	0.0022308	-0.0073897	0.0062227	-0.0068128
$\Delta\mathrm{M1}_{t-7}$	0.0335082	0.0042080	-0.0961898	-0.0054966	0.0271644	-0.6414430
$\Delta\mathrm{TB3M}_{t-7}$	0.2137423	3.2611217	0.3515946	-0.0686367	0.0791775	0.8251013
$\Delta\mathrm{OIL}_{t-7}$	-0.1884533	1.2111308	-0.7829518	-0.0030955	0.0441002	3.3629463
$\Delta\mathrm{CPI}_{t-7}$	0.0001893	0.0907228	0.0024118	0.0005227	0.0001276	0.1106807
$\Delta\mathrm{DJX}_{t-8}$	-0.0606534	-0.4685161	0.0263486	-0.0156580	0.0267021	-3.3791590
$\Delta\mathrm{IP}_{t-8}$	-0.0021734	0.0365171	0.0082247	-0.0002641	-0.0113265	0.0922530
$\Delta\mathrm{M1}_{t-8}$	-0.0053589	0.1644124	0.0275635	0.0004187	-0.0069996	1.0908443
$\Delta\mathrm{TB3M}_{t-8}$	-0.0716867	-2.5632359	0.2639982	0.0447179	-0.5230630	1.5419355
$\Delta\mathrm{OIL}_{t-8}$	0.1195372	0.2768965	0.7477269	0.0246530	-0.1091532	-0.6776440
$\Delta\mathrm{CPI}_{t-8}$	-0.0010279	-0.0212286	-0.0013249	0.0013057	-0.0068717	-0.1009224
$\Delta\mathrm{DJX}_{t-9}$	-0.0485347	-0.5241826	-0.4483325	0.0244087	0.0013183	2.9809471
$\Delta\mathrm{IP}_{t-9}$	0.0001458	0.1807325	0.0946323	0.0011936	0.0137269	0.0269402
$\Delta\mathrm{M1}_{t-9}$	-0.0101846	-0.1111001	0.0649137	-0.0018994	-0.0133131	-1.0734075
$\Delta\mathrm{TB3M}_{t-9}$	-0.1930081	2.1579470	0.9132471	-0.0969781	0.4938007	-0.1654646
$\Delta\mathrm{OIL}_{t-9}$	0.0163019	-3.9961345	-2.0131380	-0.0246940	0.0213395	-6.8958596
$\Delta\mathrm{CPI}_{t-9}$	0.0033436	-0.0708094	-0.0186001	-0.0027028	0.0029643	-0.0086340
$\Delta\mathrm{DJX}_{t-10}$	-0.0414703	-0.7864426	-0.2203061	-0.0083128	-0.1388989	0.8140576
$\Delta\mathrm{IP}_{t-10}$	-0.0026529	-0.1072024	-0.0709701	0.0026455	-0.0046504	-0.1925917

$\Delta \mathrm{M1}_{t-10}$	-0.0041652	0.0192560	-0.1066362	0.0019204	0.0155255	0.6038354
$\Delta \mathrm{TB3M}_{t-10}$	-0.0715338	-4.1181587	-0.9264807	-0.1675074	-0.0689820	-3.7253267
$\Delta \mathrm{OIL}_{t-10}$	0.0947556	0.8362993	1.3447430	0.0111884	0.2790320	5.9210939
$\Delta \mathrm{CPI}_{t-10}$	0.0043291	0.0604376	0.0372138	0.0027674	0.0014605	0.1720155
$\Delta \mathrm{DJX}_{t-11}$	0.0245572	0.6866647	0.4273090	-0.0113930	0.1404369	-1.2342392
$\Delta \mathrm{IP}_{t-11}$	0.0114305	-0.0098830	-0.0395133	-0.0047928	-0.0053725	0.4466090
$\Delta \mathrm{M1}_{t-11}$	0.0186378	-0.0995707	0.0744419	0.0088701	-0.0130863	-0.5490644
$\Delta \mathrm{TB3M}_{t-11}$	0.3729669	2.0994588	0.1830286	0.1518513	-0.0316543	6.2856308
$\Delta \mathrm{OIL}_{t-11}$	-0.2032017	1.6285527	0.6473651	-0.0181108	-0.1778579	-0.5462967
$\Delta \mathrm{CPI}_{t-11}$	-0.0076759	-0.0394912	-0.0078872	-0.0018157	-0.0003750	-0.1249999
$\Delta \mathrm{DJX}_{t-12}$	0.0039194	0.2608196	-0.2170435	-0.0025740	-0.0886585	2.1883986
$\Delta \mathrm{IP}_{t-12}$	-0.0144721	0.0235220	0.0263560	0.0025174	0.0019575	-0.1708836
$\Delta \mathrm{M1}_{t-12}$	-0.0134367	0.0374211	-0.0642185	-0.0074084	0.0134102	0.1599982
$\Delta \mathrm{TB3M}_{t-12}$	-0.1943245	0.1439329	-0.3838205	-0.0038007	-0.1420269	-2.2827826
$\Delta \mathrm{OIL}_{t-12}$	0.1020333	0.1501804	-0.6866785	0.0065312	-0.0889896	1.1830785
$\Delta \mathrm{CPI}_{t-12}$	0.0007383	0.0230952	-0.0121128	0.0005568	-0.0009787	0.0027245

B.2 Impulse Response Functions

Table B.3: Response of DJX to an impulse of DJX and confidence bands

Period	DJX	Lower	Upper
1	0.0335566	0.0261048	0.0317865
2	0.0304241	0.0214031	0.0308910
3	0.0274562	0.0174220	0.0289765
4	0.0257116	0.0145414	0.0281682
5	0.0241596	0.0128655	0.0277845
6	0.0239539	0.0124097	0.0285683
7	0.0190586	0.0067158	0.0238168
8	0.0217930	0.0106942	0.0267594
9	0.0234915	0.0121478	0.0282972
10	0.0224248	0.0118661	0.0282656
11	0.0218760	0.0100347	0.0279957
12	0.0229560	0.0113121	0.0303047
13	0.0246289	0.0118500	0.0312572
14	0.0235090	0.0094040	0.0291397
15	0.0230720	0.0089364	0.0292084
16	0.0231804	0.0097840	0.0298265
17	0.0231816	0.0098284	0.0295308
18	0.0226323	0.0083286	0.0307969
19	0.0226618	0.0087525	0.0301505
20	0.0236649	0.0098210	0.0310279
21	0.0241216	0.0098171	0.0314293
22	0.0240765	0.0099903	0.0312266
23	0.0245622	0.0098977	0.0329190
24	0.0256009	0.0094953	0.0335900
25	0.0259732	0.0097436	0.0339133
26	0.0258907	0.0100821	0.0336195

27	0.0264324	0.0103352	0.0342567
28	0.0268616	0.0106816	0.0343443
29	0.0268427	0.0111172	0.0342179
30	0.0269671	0.0114897	0.0342224
31	0.0273038	0.0113055	0.0346241
32	0.0275295	0.0114818	0.0354725
33	0.0275178	0.0114574	0.0352733
34	0.0275545	0.0113639	0.0354516
35	0.0276624	0.0114124	0.0361795
36	0.0277098	0.0115628	0.0361980
37	0.0275370	0.0112068	0.0358933

Table B.4: Response of DJX to an impulse of IP and confidence bands

Period	DJX	Lower	Upper
1	0.0004983	−0.0033819	0.0046448
2	0.0075872	0.0009575	0.0115161
3	0.0108070	0.0034780	0.0165895
4	0.0108517	0.0022776	0.0169933
5	0.0113482	−0.0003080	0.0181180
6	0.0173624	0.0059612	0.0236347
7	0.0158495	0.0046302	0.0222071
8	0.0195541	0.0057700	0.0254454
9	0.0200075	0.0060658	0.0260417
10	0.0181832	0.0045918	0.0240382
11	0.0161176	0.0034888	0.0221743
12	0.0153591	0.0002101	0.0230780
13	0.0165996	0.0002377	0.0245699
14	0.0160208	0.0001482	0.0240044
15	0.0164065	−0.0012662	0.0256952
16	0.0152725	−0.0025077	0.0256209

17	0.0152299	−0.0018121	0.0270601
18	0.0153673	−0.0009903	0.0275583
19	0.0145138	−0.0015626	0.0270130
20	0.0150459	−0.0010084	0.0276871
21	0.0156981	−0.0008254	0.0289236
22	0.0155657	−0.0028760	0.0288634
23	0.0155739	−0.0029923	0.0290862
24	0.0157331	−0.0030826	0.0299482
25	0.0161339	−0.0025415	0.0299695
26	0.0161385	−0.0020780	0.0302082
27	0.0165100	−0.0017529	0.0306937
28	0.0165761	−0.0016524	0.0310153
29	0.0164015	−0.0021333	0.0312630
30	0.0161841	−0.0031976	0.0309152
31	0.0155946	−0.0035708	0.0303957
32	0.0151357	−0.0042939	0.0299565
33	0.0149392	−0.0046153	0.0300155
34	0.0145274	−0.0046654	0.0296021
35	0.0142549	−0.0052378	0.0291119
36	0.0140932	−0.0051092	0.0289725
37	0.0136192	−0.0052847	0.0280423

Table B.5: Response of DJX to an impulse of M1 and confidence bands

Period	DJX	Lower	Upper
1	−0.0004354	−0.0035738	0.0038497
2	−0.0019338	−0.0060318	0.0038160
3	−0.0012869	−0.0057519	0.0060012
4	0.0001961	−0.0062446	0.0074073
5	0.0018468	−0.0051678	0.0080908
6	0.0080884	−0.0001886	0.0143489
7	0.0088202	0.0008145	0.0150041
8	0.0104171	0.0009189	0.0174097
9	0.0125788	0.0028919	0.0207850
10	0.0097282	−0.0002721	0.0187247
11	0.0092367	−0.0022206	0.0197271
12	0.0097110	−0.0019606	0.0220245
13	0.0091033	−0.0027111	0.0210656
14	0.0099386	−0.0023142	0.0208394
15	0.0112667	−0.0018354	0.0231963
16	0.0109617	−0.0026707	0.0229777
17	0.0119853	−0.0029680	0.0237311
18	0.0121765	−0.0027635	0.0239094
19	0.0109139	−0.0047747	0.0240718
20	0.0112872	−0.0046883	0.0237342
21	0.0115210	−0.0042233	0.0240959
22	0.0111397	−0.0042824	0.0233659
23	0.0118048	−0.0038813	0.0246664
24	0.0118272	−0.0040740	0.0255174
25	0.0119544	−0.0043397	0.0255720
26	0.0127197	−0.0042454	0.0269318
27	0.0128041	−0.0048771	0.0275227
28	0.0130734	−0.0054212	0.0282380

29	0.0135329	−0.0054947	0.0290932
30	0.0136386	−0.0059509	0.0293959
31	0.0136199	−0.0064601	0.0295366
32	0.0139611	−0.0061468	0.0300999
33	0.0140578	−0.0061693	0.0302127
34	0.0141735	−0.0064575	0.0303463
35	0.0145091	−0.0063506	0.0310066
36	0.0147297	−0.0061479	0.0314028
37	0.0147950	−0.0060655	0.0316912

Table B.6: Response of DJX to an impulse of TB3M and confidence bands

Period	DJX	Lower	Upper
1	−0.0006712	−0.0037529	0.0031768
2	−0.0069103	−0.0108738	−0.0018679
3	−0.0088414	−0.0135418	−0.0028679
4	−0.0038982	−0.0094091	0.0026465
5	−0.0045947	−0.0115983	0.0027172
6	−0.0045290	−0.0115558	0.0039019
7	−0.0074714	−0.0145019	0.0022936
8	−0.0042885	−0.0120200	0.0051830
9	−0.0037538	−0.0116420	0.0054080
10	−0.0018282	−0.0097058	0.0070537
11	0.0016008	−0.0079712	0.0112339
12	0.0073545	−0.0019829	0.0171210
13	0.0088836	−0.0013492	0.0193885
14	0.0084291	−0.0020689	0.0179829
15	0.0087012	−0.0021609	0.0188837
16	0.0100181	−0.0032010	0.0205270
17	0.0096032	−0.0034416	0.0193445
18	0.0104908	−0.0028626	0.0214360

19	0.0128584	−0.0015024	0.0245146
20	0.0135873	−0.0011997	0.0248014
21	0.0131577	−0.0025421	0.0248016
22	0.0130020	−0.0021036	0.0255692
23	0.0128604	−0.0030433	0.0248095
24	0.0120220	−0.0037975	0.0240994
25	0.0113615	−0.0048946	0.0236304
26	0.0108948	−0.0053129	0.0238079
27	0.0113012	−0.0056907	0.0247880
28	0.0110705	−0.0057248	0.0241491
29	0.0103257	−0.0069286	0.0241561
30	0.0099623	−0.0084894	0.0246673
31	0.0099013	−0.0090653	0.0248032
32	0.0093390	−0.0099850	0.0241276
33	0.0088529	−0.0103834	0.0240541
34	0.0089653	−0.0105173	0.0243321
35	0.0088457	−0.0110439	0.0243201
36	0.0084604	−0.0119242	0.0240657
37	0.0082493	−0.0122326	0.0240632

Table B.7: Response of DJX to an impulse of OIL and confidence bands

Period	DJX	Lower	Upper
1	−0.0030488	−0.0071445	0.0011723
2	−0.0083212	−0.0131297	−0.0026015
3	−0.0067876	−0.0131103	0.0002276
4	−0.0044049	−0.0108640	0.0038135
5	−0.0062102	−0.0127865	0.0040616
6	−0.0079707	−0.0148089	0.0015533
7	−0.0085404	−0.0150088	0.0016787
8	−0.0099676	−0.0158379	0.0015528

9	−0.0124815	−0.0189244	0.0011508
10	−0.0129929	−0.0205180	0.0018992
11	−0.0170558	−0.0258505	−0.0010636
12	−0.0183326	−0.0274191	−0.0001127
13	−0.0197898	−0.0299773	−0.0018388
14	−0.0218773	−0.0315773	−0.0032056
15	−0.0205954	−0.0298531	−0.0023460
16	−0.0226945	−0.0308461	−0.0040597
17	−0.0226256	−0.0316685	−0.0047664
18	−0.0222908	−0.0310414	−0.0027780
19	−0.0217925	−0.0305521	−0.0014150
20	−0.0212621	−0.0298898	−0.0000375
21	−0.0209792	−0.0309936	−0.0004908
22	−0.0193083	−0.0306703	0.0007355
23	−0.0196976	−0.0315754	0.0008937
24	−0.0203146	−0.0327209	0.0001914
25	−0.0200220	−0.0317688	0.0009817
26	−0.0205355	−0.0321396	0.0004097
27	−0.0206681	−0.0320576	−0.0001902
28	−0.0204201	−0.0315588	−0.0000028
29	−0.0203197	−0.0319916	0.0001735
30	−0.0198480	−0.0315107	0.0008340
31	−0.0200893	−0.0326206	0.0009971
32	−0.0202213	−0.0336892	0.0010338
33	−0.0203883	−0.0340492	0.0014602
34	−0.0209883	−0.0353088	0.0012249
35	−0.0215463	−0.0357773	0.0010832
36	−0.0218595	−0.0363798	0.0006391
37	−0.0220886	−0.0366821	0.0006861

Table B.8: Response of DJX to an impulse of CPI and confidence bands

Period	DJX	Lower	Upper
1	−0.0013868	−0.0040205	0.0020083
2	−0.0016384	−0.0060456	0.0041751
3	−0.0038982	−0.0082705	0.0026299
4	−0.0029603	−0.0076650	0.0030934
5	−0.0057161	−0.0111951	0.0013696
6	−0.0054531	−0.0116345	0.0028098
7	−0.0084944	−0.0150481	0.0016874
8	−0.0089691	−0.0165280	0.0020780
9	−0.0132331	−0.0188574	0.0001165
10	−0.0150708	−0.0210405	−0.0005149
11	−0.0126039	−0.0199456	0.0008859
12	−0.0126115	−0.0198128	0.0006543
13	−0.0130283	−0.0196572	−0.0001987
14	−0.0111984	−0.0178408	0.0022843
15	−0.0099205	−0.0160840	0.0037719
16	−0.0102629	−0.0176287	0.0039592
17	−0.0089966	−0.0186402	0.0053357
18	−0.0092176	−0.0174614	0.0052002
19	−0.0094265	−0.0178956	0.0064225
20	−0.0093902	−0.0189180	0.0071129
21	−0.0088109	−0.0191231	0.0078814
22	−0.0086266	−0.0189741	0.0074529
23	−0.0084227	−0.0195064	0.0066178
24	−0.0083170	−0.0193724	0.0070937
25	−0.0078436	−0.0193248	0.0082727
26	−0.0076489	−0.0189353	0.0084134
27	−0.0080157	−0.0187952	0.0089828
28	−0.0083785	−0.0189065	0.0094438

29	-0.0085390	-0.0192888	0.0101726
30	-0.0089505	-0.0196474	0.0100605
31	-0.0093475	-0.0199569	0.0100285
32	-0.0094204	-0.0200226	0.0095359
33	-0.0096129	-0.0200710	0.0096299
34	-0.0096257	-0.0205065	0.0096595
35	-0.0096217	-0.0209719	0.0096579
36	-0.0094509	-0.0210566	0.0097332
37	-0.0093759	-0.0209656	0.0093154

B.3 Forecast Error Variance Decomposition

Table B.9: Forecast error variance decomposition of DJX

Period	DJX	IP	M1	TB3M	OIL	CPI
1	1.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
2	0.9445513	0.0208773	0.0009871	0.0203348	0.0131655	0.0000840
3	0.8978678	0.0451947	0.0013637	0.0382901	0.0141367	0.0031471
4	0.8868073	0.0608969	0.0011335	0.0345741	0.0121118	0.0044763
5	0.8681465	0.0730866	0.0010075	0.0337374	0.0127346	0.0112874
6	0.8214568	0.1057689	0.0084226	0.0314832	0.0152517	0.0176168
7	0.7740594	0.1253686	0.0147448	0.0360759	0.0185149	0.0312364
8	0.7282885	0.1520545	0.0220209	0.0332038	0.0223606	0.0420716
9	0.6820448	0.1689994	0.0300671	0.0294550	0.0283552	0.0610786
10	0.6506350	0.1778529	0.0314152	0.0260293	0.0340335	0.0800340
11	0.6280089	0.1804389	0.0326426	0.0234115	0.0466820	0.0888162
12	0.6074043	0.1796142	0.0338653	0.0249222	0.0590005	0.0951936
13	0.5889316	0.1796064	0.0341978	0.0273015	0.0705306	0.0994321
14	0.5719206	0.1791438	0.0362134	0.0286955	0.0836685	0.1003581
15	0.5572544	0.1803644	0.0402812	0.0301911	0.0919018	0.1000071
16	0.5434438	0.1790039	0.0438275	0.0323097	0.1017726	0.0996425
17	0.5315105	0.1780703	0.0490217	0.0336913	0.1096205	0.0980857
18	0.5197589	0.1776740	0.0542213	0.0355706	0.1158770	0.0968983
19	0.5096722	0.1765541	0.0578141	0.0392833	0.1207923	0.0958840
20	0.5008624	0.1758301	0.0615605	0.0431229	0.1239902	0.0946339
21	0.4930835	0.1759787	0.0655469	0.0461462	0.1262830	0.0929617
22	0.4868869	0.1765773	0.0693381	0.0488342	0.1268334	0.0915301
23	0.4810666	0.1768701	0.0737173	0.0510474	0.1273626	0.0899360
24	0.4764437	0.1769990	0.0780487	0.0523589	0.1279950	0.0881547
25	0.4724245	0.1775394	0.0826038	0.0531199	0.1281505	0.0861618
26	0.4680540	0.1779022	0.0878570	0.0535114	0.1285352	0.0841402

27	0.4638211	0.1783447	0.0928841	0.0540511	0.1286950	0.0822041
28	0.4599179	0.1787551	0.0980250	0.0544239	0.1284395	0.0804387
29	0.4561057	0.1790604	0.1035213	0.0544402	0.1280469	0.0788254
30	0.4526017	0.1793138	0.1089965	0.0543593	0.1272905	0.0774382
31	0.4494825	0.1791190	0.1143216	0.0542738	0.1265974	0.0762056
32	0.4465992	0.1786460	0.1198750	0.0539871	0.1258829	0.0750098
33	0.4437471	0.1781575	0.1254183	0.0535720	0.1252159	0.0738892
34	0.4408529	0.1774279	0.1309489	0.0532322	0.1247962	0.0727419
35	0.4378792	0.1765343	0.1366177	0.0528717	0.1245361	0.0715609
36	0.4348829	0.1756361	0.1424211	0.0524222	0.1243265	0.0703112

Table B.10: Forecast error variance decomposition of IP

Period	DJX	IP	M1	TB3M	OIL	CPI
1	0.0001045	0.9998955	0.0000000	0.0000000	0.0000000	0.0000000
2	0.0011026	0.9834095	0.0008910	0.0076393	0.0020905	0.0048671
3	0.0077800	0.9615155	0.0011674	0.0194409	0.0035919	0.0065042
4	0.0452482	0.9257982	0.0009448	0.0149884	0.0023148	0.0107055
5	0.0690368	0.9044953	0.0011176	0.0101863	0.0015418	0.0136222
6	0.0840939	0.8930846	0.0022895	0.0079878	0.0011131	0.0114311
7	0.1001981	0.8776766	0.0026311	0.0093876	0.0016941	0.0084125
8	0.1141207	0.8643289	0.0031731	0.0089509	0.0025902	0.0068364
9	0.1253555	0.8495346	0.0044307	0.0082258	0.0054958	0.0069576
10	0.1302846	0.8378511	0.0077971	0.0083144	0.0073442	0.0084085
11	0.1328195	0.8229903	0.0108202	0.0099717	0.0101186	0.0132797
12	0.1385166	0.8041128	0.0144184	0.0109107	0.0147130	0.0173285
13	0.1435005	0.7873990	0.0176424	0.0119055	0.0184368	0.0211158
14	0.1475095	0.7707005	0.0212439	0.0139279	0.0230652	0.0235530
15	0.1505406	0.7534374	0.0251458	0.0168307	0.0283591	0.0256865
16	0.1523565	0.7380613	0.0293232	0.0200616	0.0329119	0.0272856
17	0.1535157	0.7230101	0.0337557	0.0235650	0.0373466	0.0288068

18	0.1542049	0.7092479	0.0384271	0.0267142	0.0418392	0.0295667
19	0.1543943	0.6962021	0.0435985	0.0295116	0.0462198	0.0300737
20	0.1538366	0.6843868	0.0488843	0.0324457	0.0500250	0.0304216
21	0.1532981	0.6732801	0.0543775	0.0351259	0.0532422	0.0306762
22	0.1529478	0.6622362	0.0594481	0.0378993	0.0568467	0.0306219
23	0.1528585	0.6517495	0.0645852	0.0406723	0.0598594	0.0302751
24	0.1528304	0.6421074	0.0699019	0.0430502	0.0623881	0.0297220
25	0.1530177	0.6330591	0.0752972	0.0451141	0.0644717	0.0290401
26	0.1533601	0.6245430	0.0806489	0.0469494	0.0661531	0.0283454
27	0.1538560	0.6165277	0.0862515	0.0482016	0.0676045	0.0275586
28	0.1544157	0.6089972	0.0919639	0.0490707	0.0688258	0.0267267
29	0.1549925	0.6018026	0.0978027	0.0496876	0.0698196	0.0258950
30	0.1556294	0.5948501	0.1036779	0.0500499	0.0706610	0.0251317
31	0.1562374	0.5882832	0.1096622	0.0501555	0.0712817	0.0243800
32	0.1568341	0.5819326	0.1156688	0.0500823	0.0718136	0.0236686
33	0.1574812	0.5757124	0.1216937	0.0498595	0.0722716	0.0229816
34	0.1581784	0.5695985	0.1277577	0.0494942	0.0726439	0.0223273
35	0.1588772	0.5636151	0.1338218	0.0490207	0.0729763	0.0216889
36	0.1595579	0.5577286	0.1398826	0.0484574	0.0732920	0.0210816

Table B.11: Forecast error variance decomposition of M1

Period	DJX	IP	M1	TB3M	OIL	CPI
1	0.0007596	0.0163042	0.9829362	0.0000000	0.0000000	0.0000000
2	0.0003841	0.0149792	0.9788594	0.0000384	0.0002724	0.0054664
3	0.0008856	0.0240058	0.9635851	0.0000262	0.0080423	0.0034549
4	0.0023052	0.0637447	0.9142969	0.0028053	0.0105363	0.0063117
5	0.0019206	0.1028648	0.8657460	0.0037233	0.0085197	0.0172257
6	0.0029773	0.1267385	0.8311190	0.0046508	0.0064075	0.0281070
7	0.0051710	0.1472293	0.8006195	0.0049063	0.0049838	0.0370900
8	0.0044290	0.1652714	0.7811907	0.0051160	0.0044490	0.0395438

9	0.0037229	0.1716785	0.7620880	0.0090732	0.0050265	0.0484110
10	0.0031459	0.1778562	0.7428478	0.0190181	0.0048894	0.0522426
11	0.0031652	0.1814857	0.7214288	0.0302893	0.0059818	0.0576492
12	0.0029212	0.1835417	0.7080305	0.0369591	0.0060522	0.0624953
13	0.0027054	0.1841168	0.6958511	0.0441429	0.0053124	0.0678714
14	0.0023485	0.1838121	0.6838890	0.0515107	0.0049029	0.0735369
15	0.0020608	0.1825304	0.6743949	0.0571383	0.0045796	0.0792961
16	0.0018199	0.1799561	0.6661247	0.0636293	0.0048368	0.0836332
17	0.0016061	0.1776117	0.6586937	0.0703365	0.0051885	0.0865635
18	0.0014541	0.1768340	0.6498398	0.0770929	0.0055315	0.0892477
19	0.0012979	0.1757406	0.6425681	0.0832411	0.0058766	0.0912758
20	0.0011621	0.1750426	0.6349928	0.0894195	0.0059418	0.0934412
21	0.0010481	0.1737383	0.6296663	0.0946505	0.0059168	0.0949799
22	0.0009519	0.1715624	0.6255948	0.0997370	0.0058498	0.0963041
23	0.0008676	0.1695410	0.6221972	0.1040437	0.0056899	0.0976605
24	0.0007934	0.1673720	0.6193284	0.1078330	0.0054601	0.0992131
25	0.0007282	0.1653765	0.6164008	0.1116240	0.0052502	0.1006204
26	0.0006713	0.1634887	0.6137048	0.1148715	0.0051291	0.1021346
27	0.0006215	0.1617963	0.6114745	0.1176178	0.0050050	0.1034849
28	0.0005774	0.1601140	0.6096108	0.1200202	0.0049070	0.1047706
29	0.0005385	0.1585384	0.6079074	0.1221553	0.0048382	0.1060221
30	0.0005037	0.1568249	0.6067268	0.1239330	0.0047610	0.1072505
31	0.0004725	0.1550694	0.6058296	0.1256362	0.0046890	0.1083033
32	0.0004445	0.1532259	0.6052775	0.1271270	0.0046040	0.1093211
33	0.0004193	0.1512708	0.6049750	0.1284504	0.0045153	0.1103693
34	0.0003961	0.1493215	0.6048275	0.1295794	0.0044203	0.1114551
35	0.0003749	0.1473905	0.6048817	0.1304576	0.0043299	0.1125654
36	0.0003555	0.1454153	0.6051669	0.1311596	0.0042575	0.1136452

Table B.12: Forecast error variance decomposition of TB3M

Period	DJX	IP	M1	TB3M	OIL	CPI
1	0.0006935	0.0098816	0.0027708	0.9866541	0.0000000	0.0000000
2	0.0019229	0.0307730	0.0020308	0.9628730	0.0019369	0.0004634
3	0.0140167	0.0508026	0.0020854	0.9277530	0.0050258	0.0003164
4	0.0164154	0.0795617	0.0016813	0.8951609	0.0064593	0.0007215
5	0.0159189	0.0957751	0.0013886	0.8802490	0.0053443	0.0013241
6	0.0151813	0.1033616	0.0011364	0.8743607	0.0042203	0.0017397
7	0.0140668	0.1057743	0.0008979	0.8726108	0.0037372	0.0029131
8	0.0129097	0.1110092	0.0007565	0.8672390	0.0030456	0.0050400
9	0.0123562	0.1129159	0.0006818	0.8631901	0.0032473	0.0076086
10	0.0121387	0.1166066	0.0008076	0.8577109	0.0029947	0.0097416
11	0.0124347	0.1184802	0.0009763	0.8524120	0.0024936	0.0132031
12	0.0134973	0.1177546	0.0012048	0.8491353	0.0021945	0.0162135
13	0.0146489	0.1160950	0.0015124	0.8470025	0.0019771	0.0187641
14	0.0154732	0.1138558	0.0018777	0.8459534	0.0018501	0.0209899
15	0.0162079	0.1116250	0.0025108	0.8452350	0.0017107	0.0227107
16	0.0169720	0.1097984	0.0033720	0.8443558	0.0015387	0.0239632
17	0.0175893	0.1081993	0.0044392	0.8430614	0.0014037	0.0253072
18	0.0179734	0.1066182	0.0059044	0.8415571	0.0013047	0.0266421
19	0.0183235	0.1052991	0.0077540	0.8396017	0.0012064	0.0278153
20	0.0187102	0.1032856	0.0094030	0.8383631	0.0011194	0.0291188
21	0.0191565	0.1006506	0.0110250	0.8379081	0.0010445	0.0302153
22	0.0196293	0.0979158	0.0125918	0.8378097	0.0009822	0.0310711
23	0.0201325	0.0950355	0.0140798	0.8380550	0.0009316	0.0317655
24	0.0206563	0.0921211	0.0156675	0.8382548	0.0008826	0.0324177
25	0.0211988	0.0894136	0.0174636	0.8381802	0.0008379	0.0329059
26	0.0217307	0.0868374	0.0194090	0.8379539	0.0007993	0.0332696
27	0.0222467	0.0843482	0.0215423	0.8375922	0.0007710	0.0334997
28	0.0227692	0.0819783	0.0237957	0.8370567	0.0007508	0.0336492

29	0.0232665	0.0796686	0.0260324	0.8366144	0.0007341	0.0336840
30	0.0237577	0.0774041	0.0282606	0.8362163	0.0007191	0.0336422
31	0.0242796	0.0752246	0.0304146	0.8358204	0.0007027	0.0335581
32	0.0248288	0.0731307	0.0325139	0.8353950	0.0006829	0.0334488
33	0.0253804	0.0711485	0.0346213	0.8348643	0.0006632	0.0333222
34	0.0259411	0.0692879	0.0367320	0.8341977	0.0006450	0.0331963
35	0.0265168	0.0675374	0.0388290	0.8334298	0.0006291	0.0330579
36	0.0271073	0.0659005	0.0409669	0.8325272	0.0006176	0.0328804

Table B.13: Forecast error variance decomposition of OIL

Period	DJX	IP	M1	TB3M	OIL	CPI
1	0.0077764	0.0008158	0.0002229	0.0010953	0.9900895	0.0000000
2	0.0054781	0.0059328	0.0010191	0.0006965	0.9795392	0.0073343
3	0.0033431	0.0100651	0.0022613	0.0018237	0.9747739	0.0077331
4	0.0027501	0.0219241	0.0045708	0.0014939	0.9596574	0.0096038
5	0.0027369	0.0315784	0.0041372	0.0011890	0.9505825	0.0097760
6	0.0027968	0.0476096	0.0037929	0.0012987	0.9301505	0.0143515
7	0.0029188	0.0654476	0.0033446	0.0019022	0.9039623	0.0224245
8	0.0027449	0.0862387	0.0035393	0.0027051	0.8760759	0.0286961
9	0.0028276	0.1088841	0.0048712	0.0025113	0.8489750	0.0319308
10	0.0026399	0.1215796	0.0047307	0.0027702	0.8298191	0.0384606
11	0.0025277	0.1261996	0.0043497	0.0041802	0.8169479	0.0457949
12	0.0023507	0.1271076	0.0040097	0.0052123	0.8068312	0.0544885
13	0.0021375	0.1247410	0.0036540	0.0073629	0.7985865	0.0635181
14	0.0019548	0.1206590	0.0033437	0.0109039	0.7904854	0.0726532
15	0.0018216	0.1174368	0.0031155	0.0134474	0.7845915	0.0795871
16	0.0017768	0.1151428	0.0029752	0.0153191	0.7800924	0.0846938
17	0.0019773	0.1146206	0.0030464	0.0170170	0.7740391	0.0892997
18	0.0022453	0.1151454	0.0035443	0.0191088	0.7669364	0.0930198
19	0.0026065	0.1160172	0.0041227	0.0215168	0.7606252	0.0951116

20	0.0029793	0.1159162	0.0045781	0.0240737	0.7564161	0.0960366
21	0.0032775	0.1146902	0.0046619	0.0269386	0.7537461	0.0966856
22	0.0035326	0.1124872	0.0045900	0.0298411	0.7528970	0.0966521
23	0.0037929	0.1098793	0.0044733	0.0324081	0.7529616	0.0964848
24	0.0040538	0.1070615	0.0043198	0.0344758	0.7536540	0.0964351
25	0.0042043	0.1045199	0.0041888	0.0361748	0.7543447	0.0965675
26	0.0043273	0.1026077	0.0041679	0.0371665	0.7552698	0.0964607
27	0.0044737	0.1014335	0.0042627	0.0376710	0.7557317	0.0964274
28	0.0046348	0.1007228	0.0044344	0.0381241	0.7556651	0.0964189
29	0.0047421	0.1001759	0.0046692	0.0385612	0.7555975	0.0962542
30	0.0047896	0.0996913	0.0048916	0.0389156	0.7558857	0.0958261
31	0.0048240	0.0989424	0.0050351	0.0392373	0.7565929	0.0953684
32	0.0048458	0.0979268	0.0051218	0.0394987	0.7576574	0.0949495
33	0.0048475	0.0967026	0.0051713	0.0396344	0.7589916	0.0946527
34	0.0048528	0.0954068	0.0051998	0.0396575	0.7603649	0.0945183
35	0.0048740	0.0941328	0.0052471	0.0395468	0.7616263	0.0945731
36	0.0048974	0.0930196	0.0053361	0.0393887	0.7626259	0.0947323

Table B.14: Forecast error variance decomposition of CPI

Period	DJX	IP	M1	TB3M	OIL	CPI
1	0.0072822	0.0224617	0.0209445	0.0071996	0.2363361	0.7057759
2	0.0023668	0.0111886	0.0162670	0.0023084	0.4657367	0.5021326
3	0.0014962	0.0113057	0.0166290	0.0030879	0.5742931	0.3931881
4	0.0013432	0.0273638	0.0180573	0.0020518	0.6474259	0.3037581
5	0.0010407	0.0414126	0.0201454	0.0026356	0.6804291	0.2543366
6	0.0008711	0.0508442	0.0245168	0.0031684	0.6995501	0.2210494
7	0.0010833	0.0588999	0.0294089	0.0055793	0.7096453	0.1953833
8	0.0024285	0.0720657	0.0367975	0.0095140	0.7069807	0.1722138
9	0.0053016	0.0949462	0.0446757	0.0088221	0.6924112	0.1538432
10	0.0064171	0.1153312	0.0517833	0.0078930	0.6778328	0.1407426

11	0.0058099	0.1336798	0.0585766	0.0077510	0.6627340	0.1314486
12	0.0051892	0.1458343	0.0664928	0.0076868	0.6521493	0.1226476
13	0.0047393	0.1554916	0.0730276	0.0081420	0.6467882	0.1118114
14	0.0045280	0.1623317	0.0789334	0.0097670	0.6421982	0.1022417
15	0.0045498	0.1676928	0.0850036	0.0120102	0.6362968	0.0944468
16	0.0049723	0.1717809	0.0918914	0.0139819	0.6288972	0.0884763
17	0.0057594	0.1763715	0.0996590	0.0164262	0.6185241	0.0832599
18	0.0066645	0.1829031	0.1088379	0.0195135	0.6041234	0.0779577
19	0.0076462	0.1902798	0.1187038	0.0227297	0.5878756	0.0727649
20	0.0086402	0.1969116	0.1280739	0.0263476	0.5715159	0.0685109
21	0.0093156	0.2018554	0.1361569	0.0306012	0.5568501	0.0652207
22	0.0096556	0.2052160	0.1430747	0.0352372	0.5441574	0.0626590
23	0.0099852	0.2078494	0.1491812	0.0396828	0.5331454	0.0601559
24	0.0103263	0.2104810	0.1547687	0.0441868	0.5225335	0.0577037
25	0.0105565	0.2127369	0.1602198	0.0485797	0.5124960	0.0554111
26	0.0106730	0.2150660	0.1661793	0.0522811	0.5023378	0.0534627
27	0.0108056	0.2176249	0.1726979	0.0550415	0.4921443	0.0516858
28	0.0109823	0.2204325	0.1793226	0.0574623	0.4819005	0.0498998
29	0.0111003	0.2232573	0.1861557	0.0597930	0.4715011	0.0481926
30	0.0111290	0.2260840	0.1928749	0.0620701	0.4611536	0.0466884
31	0.0110867	0.2285609	0.1993430	0.0642922	0.4513113	0.0454059
32	0.0109887	0.2306589	0.2054929	0.0665162	0.4421407	0.0442026
33	0.0108432	0.2324613	0.2114264	0.0686023	0.4336675	0.0429992
34	0.0106942	0.2340797	0.2172470	0.0704744	0.4257098	0.0417949
35	0.0105545	0.2354281	0.2231437	0.0721303	0.4180677	0.0406758
36	0.0104090	0.2365744	0.2291785	0.0736180	0.4105755	0.0396445

B.4 Out-of-Sample Forecast

Table B.15: Forecast of DJX and confidence bands

Period	Forecast	Observed Value	Lower	Upper
1	9.5088668	9.5057770	9.4406420	9.5770916
2	9.5248279	9.4800972	9.4275236	9.6221323
3	9.5622979	9.4746704	9.4438082	9.6807877
4	9.5886066	9.4806835	9.4550985	9.7221148
5	9.5992625	9.5368041	9.4522400	9.7462850
6	9.6210286	9.5506972	9.4590695	9.7829878
7	9.6376681	9.5873059	9.4641884	9.8111478
8	9.6473575	9.6050680	9.4602248	9.8344903
9	9.6438888	9.6234806	9.4409045	9.8468731
10	9.6582275	9.6097606	9.4411749	9.8752802
11	9.6567093	9.6485656	9.4263555	9.8870632
12	9.6639369	9.6030788	9.4199643	9.9079095
13	9.6898132	9.6244130	9.4309161	9.9487102
14	9.6956453	9.6515426	9.4228858	9.9684048
15	9.7075994	9.6857301	9.4224156	9.9927833
16	9.7196329	9.7157510	9.4218716	10.0173942
17	9.7242104	9.6613427	9.4143893	10.0340315
18	9.7336971	9.7002514	9.4124733	10.0549210
19	9.7403445	9.7085463	9.4082362	10.0724528
20	9.7497945	9.7160031	9.4068069	10.0927821
21	9.7603091	9.7241916	9.4065793	10.1140388
22	9.7697284	9.7307162	9.4062300	10.1332267
23	9.7784613	9.7149447	9.4051419	10.1517807
24	9.7880803	9.7467431	9.4047940	10.1713667
25	9.7979191	9.7434890	9.4048957	10.1909426
26	9.8049376	9.7636805	9.4022198	10.2076554

27	9.8127138	9.7885390	9.4001904	10.2252371
28	9.8234985	9.7882490	9.4011550	10.2458421
29	9.8303052	9.7506248	9.3984486	10.2621617
30	9.8373567	9.7950897	9.3962878	10.2784255
31	9.8459851	9.7943282	9.3958464	10.2961239
32	9.8521337	9.7964874	9.3930781	10.3111894
33	9.8586393	9.8050315	9.3909075	10.3263711
34	9.8649112	9.7940755	9.3886194	10.3412030
35	9.8709476	9.7866739	9.3861342	10.3557610
36	9.8772900	9.7445850	9.3840825	10.3704976

Table B.16: Forecast of IP and confidence bands

Period	Forecast	Observed Value	Lower	Upper
1	4.6088247	4.6043589	4.5994317	4.6182176
2	4.6140680	4.6070464	4.6012193	4.6269168
3	4.6164110	4.6115846	4.5996170	4.6332051
4	4.6224810	4.6137136	4.6007879	4.6441740
5	4.6277075	4.6144569	4.6009738	4.6544412
6	4.6312589	4.6185059	4.5996148	4.6629029
7	4.6364977	4.6206596	4.5993048	4.6736906
8	4.6416197	4.6204380	4.5995774	4.6836620
9	4.6458037	4.6197523	4.5988447	4.6927627
10	4.6482800	4.6216546	4.5962775	4.7002825
11	4.6540292	4.6177754	4.5970677	4.7109908
12	4.6570247	4.6254061	4.5951138	4.7189355
13	4.6603895	4.6311897	4.5938380	4.7269410
14	4.6641141	4.6313583	4.5930131	4.7352151
15	4.6676375	4.6339160	4.5920584	4.7432166
16	4.6709289	4.6365613	4.5909306	4.7509272
17	4.6739534	4.6347746	4.5897377	4.7581691

18	4.6762065	4.6425421	4.5879254	4.7644876
19	4.6792579	4.6507313	4.5870473	4.7714684
20	4.6810525	4.6526214	4.5851125	4.7769925
21	4.6827206	4.6563014	4.5832163	4.7822249
22	4.6847137	4.6607563	4.5816975	4.7877298
23	4.6870198	4.6641964	4.5806311	4.7934085
24	4.6893567	4.6645121	4.5797174	4.7989959
25	4.6911762	4.6698112	4.5784031	4.8039493
26	4.6934767	4.6713914	4.5776703	4.8092831
27	4.6954260	4.6802739	4.5767087	4.8141433
28	4.6973035	4.6813050	4.5757819	4.8188250
29	4.6991923	4.6784234	4.5749519	4.8234327
30	4.7010755	4.6769028	4.5741939	4.8279572
31	4.7026063	4.6750451	4.5731675	4.8320450
32	4.7044434	4.6733885	4.5725272	4.8363597
33	4.7060138	4.6698328	4.5716986	4.8403290
34	4.7077362	4.6698346	4.5710929	4.8443795
35	4.7092656	4.6775913	4.5703636	4.8481675
36	4.7107451	4.6784857	4.5696538	4.8518363

Table B.17: Forecast of M1 and confidence bands

Period	Forecast	Observed Value	Lower	Upper
1	7.7635834	7.7763672	7.7505222	7.7766445
2	7.7702533	7.7896616	7.7518572	7.7886493
3	7.7751813	7.7928443	7.7521592	7.7982034
4	7.7812587	7.8069812	7.7522980	7.8102195
5	7.7842540	7.8110013	7.7501398	7.8183682
6	7.7904440	7.8121354	7.7512228	7.8296653
7	7.7946346	7.8139149	7.7494285	7.8398407
8	7.8014492	7.8284364	7.7510109	7.8518874

9	7.8079631	7.8328075	7.7519885	7.8639376
10	7.8127796	7.8311805	7.7504603	7.8750990
11	7.8161273	7.8423966	7.7475804	7.8846741
12	7.8198546	7.8474109	7.7447812	7.8949279
13	7.8226406	7.8555834	7.7409998	7.9042813
14	7.8287761	7.8709678	7.7408902	7.9166620
15	7.8329670	7.8717689	7.7387172	7.9272168
16	7.8362714	7.8840116	7.7355420	7.9370008
17	7.8414630	7.8959573	7.7345034	7.9484227
18	7.8438157	7.9090854	7.7304144	7.9572170
19	7.8478653	7.9187014	7.7281470	7.9675836
20	7.8519276	7.9282979	7.7259249	7.9779302
21	7.8556536	7.9330439	7.7234879	7.9878193
22	7.8586660	7.9431437	7.7203573	7.9969746
23	7.8623882	7.9525096	7.7180846	8.0066918
24	7.8660096	7.9442084	7.7157200	8.0162993
25	7.8698953	7.9588216	7.7136402	8.0261504
26	7.8731743	7.9612307	7.7110678	8.0352808
27	7.8765185	7.9667601	7.7086807	8.0443562
28	7.8801675	7.9787565	7.7067558	8.0535793
29	7.8839252	7.9847353	7.7050525	8.0627978
30	7.8874430	8.0056340	7.7032156	8.0716704
31	7.8913243	8.0039647	7.7018580	8.0807906
32	7.8950641	8.0065675	7.7004746	8.0896536
33	7.8985517	8.0015895	7.6989040	8.0981993
34	7.9024043	8.0095625	7.6978197	8.1069890
35	7.9064458	8.0203691	7.6970443	8.1158472
36	7.9103946	8.0262682	7.6962885	8.1245006

Table B.18: Forecast of TB3M and confidence bands

Period	Forecast	Observed Value	Lower	Upper
1	0.2462298	0.11	−0.0342278	0.5266874
2	0.4805894	0.10	0.0099353	0.9512434
3	0.5501131	0.09	−0.0775088	1.1777350
4	0.5648962	0.07	−0.2402380	1.3700303
5	0.6828832	0.07	−0.2870235	1.6527900
6	0.7536096	0.10	−0.3503944	1.8576137
7	0.7407817	0.09	−0.5020029	1.9835663
8	0.7666285	0.06	−0.6287568	2.1620138
9	0.8108193	0.04	−0.7450803	2.3667188
10	0.8777745	0.05	−0.8585876	2.6141366
11	0.9598530	0.04	−0.9491539	2.8688599
12	0.9929476	0.04	−1.0810798	3.0669749
13	1.0505633	0.02	−1.1825433	3.2836700
14	1.1144013	0.05	−1.2676973	3.4964998
15	1.1308752	0.07	−1.3916814	3.6534319
16	1.1417704	0.07	−1.5209583	3.8044990
17	1.1908127	0.04	−1.6083408	3.9899662
18	1.1977010	0.05	−1.7322688	4.1276708
19	1.1849431	0.05	−1.8713445	4.2412307
20	1.1916016	0.03	−1.9865629	4.3697660
21	1.1837892	0.03	−2.1107808	4.4783593
22	1.1736894	0.04	−2.2312709	4.5786498
23	1.1669883	0.03	−2.3442076	4.6781842
24	1.1567045	0.03	−2.4574842	4.7708933
25	1.1408595	0.02	−2.5720190	4.8537381
26	1.1280544	0.02	−2.6779361	4.9340449
27	1.0945965	0.02	−2.7993439	4.9885370
28	1.0603323	0.03	−2.9171730	5.0378375

29	1.0254954	0.03	−3.0318565	5.0828473
30	0.9869142	0.02	−3.1469312	5.1207596
31	0.9478252	0.03	−3.2593440	5.1549944
32	0.9160510	0.02	−3.3614862	5.1935883
33	0.8777744	0.02	−3.4665941	5.2221429
34	0.8362810	0.02	−3.5713529	5.2439149
35	0.7943871	0.03	−3.6734540	5.2622282
36	0.7481311	0.07	−3.7769787	5.2732409

Table B.19: Forecast of OIL and confidence bands

Period	Forecast	Observed Value	Lower	Upper
1	5.3488043	5.3007145	5.2140626	5.4835461
2	5.3831782	5.2733072	5.1631053	5.6032511
3	5.3889960	5.2519069	5.1048475	5.6731444
4	5.4744225	5.2512782	5.1333402	5.8155048
5	5.4913602	5.2878124	5.1117384	5.8709820
6	5.5121322	5.3129104	5.1009627	5.9233017
7	5.5304794	5.2645020	5.0970457	5.9639130
8	5.5101733	5.2268751	5.0583036	5.9620430
9	5.4958698	5.2310017	5.0279617	5.9637779
10	5.5385250	5.2347384	5.0588834	6.0181666
11	5.5520098	5.2869025	5.0586841	6.0453355
12	5.5702286	5.3138954	5.0591548	6.0813024
13	5.5798828	5.3209591	5.0509750	6.1087907
14	5.6012933	5.2906380	5.0547494	6.1478371
15	5.6152303	5.2642434	5.0542656	6.1761950
16	5.6222540	5.2917458	5.0486969	6.1958111
17	5.6252398	5.2604078	5.0415009	6.2089786
18	5.6247997	5.2842184	5.0317274	6.2178719
19	5.6258220	5.2766851	5.0236380	6.2280061

20	5.6279044	5.2851815	5.0170815	6.2387273
21	5.6334920	5.2927016	5.0132084	6.2537755
22	5.6387119	5.3172863	5.0073887	6.2700350
23	5.6467389	5.2879640	5.0037800	6.2896978
24	5.6507891	5.2382488	4.9962825	6.3052958
25	5.6639566	5.1953983	4.9982584	6.3296548
26	5.6671381	5.0877815	4.9915405	6.3427358
27	5.6717208	4.9748008	4.9869745	6.3564672
28	5.6728983	4.7351453	4.9797534	6.3660431
29	5.6740973	4.4903203	4.9730265	6.3751681
30	5.6783065	4.6392817	4.9697125	6.3869005
31	5.6845797	4.6009613	4.9683678	6.4007917
32	5.6885261	4.6832417	4.9642274	6.4128249
33	5.6925970	4.7678841	4.9598944	6.4252996
34	5.6967131	4.7478843	4.9556423	6.4377839
35	5.7006291	4.6302530	4.9514135	6.4498447
36	5.7032117	4.4559739	4.9462083	6.4602151

Table B.20: Forecast of CPI and confidence bands

Period	Forecast	Observed Value	Lower	Upper
1	5.4420859	5.4428678	5.4386009	5.4455709
2	5.4455195	5.4456204	5.4393984	5.4516406
3	5.4489359	5.4432918	5.4407657	5.4571061
4	5.4527916	5.4431317	5.4427804	5.4628028
5	5.4553775	5.4443379	5.4437170	5.4670380
6	5.4584797	5.4501926	5.4455314	5.4714280
7	5.4609127	5.4477928	5.4469465	5.4748790
8	5.4624723	5.4453226	5.4474562	5.4774884
9	5.4641232	5.4466943	5.4481813	5.4800651
10	5.4671350	5.4492472	5.4504864	5.4837837

11	5.4710444	5.4509526	5.4537951	5.4882938
12	5.4739943	5.4528094	5.4560553	5.4919333
13	5.4769219	5.4543506	5.4582395	5.4956043
14	5.4795904	5.4549065	5.4602300	5.4989509
15	5.4824481	5.4554835	5.4624404	5.5024557
16	5.4852008	5.4582953	5.4645720	5.5058296
17	5.4878613	5.4601300	5.4666454	5.5090773
18	5.4903955	5.4610993	5.4685364	5.5122546
19	5.4929113	5.4629416	5.4703731	5.5154495
20	5.4951861	5.4648482	5.4719631	5.5184091
21	5.4979591	5.4678491	5.4740813	5.5218369
22	5.5006969	5.4695274	5.4761566	5.5252373
23	5.5033716	5.4705718	5.4781320	5.5286112
24	5.5059711	5.4697844	5.4800150	5.5319273
25	5.5087516	5.4706980	5.4820825	5.5354207
26	5.5115037	5.4712323	5.4841266	5.5388808
27	5.5141263	5.4683428	5.4860414	5.5422111
28	5.5166449	5.4650345	5.4878555	5.5454344
29	5.5190583	5.4582101	5.4895662	5.5485504
30	5.5215415	5.4603767	5.4913572	5.5517258
31	5.5241154	5.4627295	5.4932537	5.5549772
32	5.5266894	5.4637555	5.4951592	5.5582197
33	5.5291727	5.4681909	5.4969680	5.5613774
34	5.5316812	5.4713711	5.4988091	5.5645532
35	5.5340824	5.4726866	5.5005585	5.5676064
36	5.5365118	5.4719807	5.5023510	5.5706725

List of Figures

List of Tables

Bibliography

[1] Abdullah, D. A. and Hayworth, S. C. (1993): *Macroeconometrics of Stock Price Fluctuations*. Quarterly Journal of Business and Economics, 32, 46–63.

[2] Agrawalla, R. K. and Tuteja, S. K. (2007): *Causality between Stock Market Development and Economic Growth: a Case Study of India*. Journal of Management Research, 7, 158–168.

[3] Akaike, H. (1974): *A New Look at the Statistical Model Identification*. IEEE Transactions on Automatic Control, 19(6), 716–23.

[4] Allen L. and Jagtianti, J. (1997): *Risk and Market Segmentation in Financial Intermediaries' Returns*. Journal of Financial Service Research, 12, 159–173.

[5] Apergis, N. and Eleftheriou, S. (2002): *Interest Rates, Inflation, and Stock Prices: The Case of the Athens Stock Exchange*. Journal of Policy Modeling 24, 231–236.

[6] Asprem, M. (1989): *Stock Prices, Asset Portfolios and Macroeconomic Variables in Ten European Countries*. Journal of Banking and Finance, 13, 589–612.

[7] Ahtola, J. and Tiao, G. C. (1987): *Distributions of Least Squares Estimators of Autoregressive Parameters for a Process with Complex Roots on the Unit Circle*. Journal of Time Series Analysis, 8, 1–14.

[8] Banerjee, A., Dolado, J. J., Galbraith, J. W. and Hendry, D. F. (1993): *Co-integration, Error Correction and the Econometric Analysis of Non-stationary Data*. Oxford University Press, New York.

[9] Beaulieu, J. J. and Miron, J. A. (1993): *Seasonal Unit Roots in Aggregate U.S. Data*. Journal of Econometrics, 55, 305–328.

[10] Bhargava, A. (1987): *On the Specification of Regression Models in Seasonal Differences*. Mimeo. (Department of Economics, University of Pennsylvania, Philadelphia, PA). ,

[11] Bossaerts, P. (1988): *Common Nonstationary Components of Asset Pricing*. Journal of Economic Dynamics and Control, 12, 347–364.

[12] Box, G. E. P. and Jenkins, G. M. (1976): *Time Series Analysis: Forecasting and Control*. 2nd edn., Holden-Day, San Francisco.

[13] Box, G. E. P. and Pierce, D. A. (1970): *Distribution of the Residual Autocorrelations in Autoregressive-Integrated Moving-Average Time Series Models*. Journal of the American Statistical Association, 65, 1509–1526.

[14] Breusch, T. and Pagan, A. (1979): *A Simple Test of Heteroskedasticity and Random Coefficient Variation*. Econometrica 47, 1287–1294.

[15] Busetti, F. and Taylor, A. M. R. (2003): *Testing Against Stochastic Trend in the Presence of Variance Shifts*. Journal of Business and Economic Statistics, 21, 510–531.

[16] Catani, P. (2013): *Misspecification and Bootstrap Tests in Multivariate Time Series Models with Conditional Heteroskedasticity*. Ekonomi och Samhälle / Economics and Society, 262. Hanken School of Economics, Helsinki.

[17] Cavaliere, G. and Taylor, A. M. R. (2006): *Testing the Null of Cointegration in the Presence of Variance Breaks*. Journal of Time Series Analysis, 27, 619–636.

[18] Cavaliere, G. and Taylor, A. M. R. (2007): *Testing for Unit Roots in Time Series Models with Nonstationary Volatility*. Journal of Econometrics, 140, 919–947.

[19] Cavaliere, G., Rahbek, A. and Taylor, A. M. R. (2010a): *Cointegration Rank Testing under Conditional Heteroskedasticity*. Economic Theory, 26, 1719–1760.

[20] Cavaliere, G., Rahbek, A. and Taylor, A. M. R. (2010b): *Testing for Cointegration in Vector Autoregressions with Nonstationary Volatility*. Journal of Econometrics, 158, 7–24.

[21] Chaudhuri, K. and Smile, S. (2004): *Stock Market and Aggregate Economic Activity: Evidence from Australia*, Applied Financial Economics, 14, 121–129.

[22] Chen, N. F., Roll, R. and Ross, S. (1986): *Economic Forces and the Stock Market*. Journal of Business, 59(3), 83–403.

[23] Cheng, A. C. S. (1995): *The UK Stock Market and Economic Factors: a New Approach*. Journal of Business Finance and Accounting, 22, 129–142.

[24] Cheung Y. W. and Lai, K. S. (1993): *Finite Sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration*. Oxford Bulletin of Economics and Statistics, 55, 313–328.

[25] Cheung, Y. and Ng, L. (1998): *International Evidence on the Stock Market and Aggregate Economic Activity*, Journal of Empirical Finance, 5, 281–296.

[26] Cleveland, W. S. (1979): *Robust Locally Weighted Regression and Smoothing Scatterplots*. Journal of the American Statistical Association, 74, 829–836.

[27] Cologni, A. and Manera, M. (2008): *Oil Prices, Inflation and Interest Rates in a Structural Cointegrated VAR Model for the G-7 Countries*. Energy Economics, 30, 856–888.

[28] Cooper, R. L. (1972): *The Predictive Performance of Quarterly Econometric Models of the United States*. In Econometric Models of Cyclical Behavior (B.G. Hickman, ed.) New York: National Bureau of Economic Research.

[29] Culter, D. M., Poterba, J. M. and Summers, L. H. (1989): *What Moves Stock Prices?* Journal of Portfolio Management, 15, 4–12.

[30] Darrat, A. F. (1990): *Stock Returns, Money and Fiscal Policy*. Journal of Financial and Quantitative Analysis, 25, 387–398.

[31] Darby, M. R. (1982): *The Price of Oil and World Inflation and Recession*. American Economic Review, 72, 738–751.

[32] Dickey, D. A. and Fuller, W. A. (1979): *Distribution of the Estimators for Autoregressive Time Series with a Unit Root*. Journal of the American Statistical Association, 74, 427–431.

[33] Dickey, D. A. and Fuller, W. A. (1981): *Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root*. Econometrica, 49, 1057–1072.

[34] Dickey, D. A., Hasza, H. P. and Fuller W. A. (1984): *Testing for Unit Roots in Seasonal Times Series*. Journal of the American Statistical Association, 79, 355–367.

[35] Dritsaki, C. and Adamopoulos, A. A. (2005): *Causal Relationship and Macroeconomic Activity: Empirical Results from European Union*. American Journal of Applied Science, 2, 504–507.

[36] Enders, W. (1995): *Applied Econometric Time Series*. John Wiley & Sons, New York.

[37] Engle R. F. (1982): *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*. Econometrica, 50, 987–1007.

[38] Engle, R. F. and Yoo, B. S. (1987): *Forecasting and Testing in Co-integrated Systems*. Journal of Econometrics, 35, 143–159.

[39] Engle, R. F. and Granger, C. W. J. (1987): *Cointegration and Error Correction: Representation, Estimation and Testing*. Econometrica, 55, 251–276.

[40] Fama, E. F. (1970): *Efficient Capital Markets: a Review of Theory and Empirical Work*. Journal of Finance, 25, 383–417.

[41] Fama, E. F. (1975): *Short-Term Interest Rates as Predictors of Inflation*. American Economic Review, 269–282.

[42] Fama, E. F. (1976): *Inflation Uncertainty and the Expected Return on Treasury Bills*. Journal of Political Economy, 427–448.

[43] Fama, E. F. (1981): *Stock Returns, Real Activity, Inflation and Money*. The American Economic Review, 71(4), 45–565.

[44] Fama, E. F. (1990): *Stock Returns, Expected Returns, and Real Activity*. Journal of Finance, 45, 1089–1108.

[45] Fama, E. F. and Gibbons, M. (1982): *Inflation, Real Returns and Capital Investment*. Journal of Monetary Economics, 397–424.

[46] Fama, E. F. and Schwert, W. G. (1977): *Asset Returns and Inflation*. Journal of Financial Economics 5, 115–146.

[47] Ferson, W. E. and Harvey, C. R. (1991): *The Variation of Economic Risk Premiums*. Journal of Political Economy, 99, 385–415.

[48] Ferson, W. E and Harvey, C. R. (1993): *The Risk and Predictability of International Equity Returns*. Review of Financial Studies, 6, 527–566.

[49] Fifield, S. Power, D. and Sinclair, C. (2000): *A Study of Whether Macroeconomic Factors Influence Emerging Market Share Seturns*. Global Economy Quarterly, 1(1), 315–335.

[50] Flannery, M. J. and James, C. M. (1984): *The Effect of Interest Rate Changes on the Common Stock Returns of Financial Institutions*. Journal of Finance, 1141–1153.

[51] Franses, P. H. and Hobijn, B. (1997): *Critical Values for Unit Root Tests in Seasonal Time Series*. Journal of Applied Statistics, 24, 25–48.

[52] Franses, P. H. (1991): *Seasonality, Non-stationarity and the Forecasting of Monthly Time Series*. International Journal of Forecasting, 7(2), 199–208.

[53] French, K. R., Ruback, R. C., Schwert, G. W. (1983): *Effects of Nominal Contracting on Stock Returns*. Journal of Political Economy, 70–96.

[54] Friedman, M. and Schwart., A. J. (1963): *Money and Business Cycles*. Review of Economics and Statistics, 45(1), 485.

[55] Fuller, W. A. (1976): *Introduction to Statistical Time Series*. John Wiley & Sons, New York.

[56] Geske, R. and Roll R. (1983): *The Fiscal and Monetary Linkage between Stock Returns and Inflation*. Journal of Finance, 38, 7–33.

[57] Gjerde, O. and Saettem, F. (1999): *Causal Relation among Stock Returns and Macroeconomic Variables in a Small, Open Economy*. Journal of International Financial Markets, Institutions and Money, 9, 61–74.

[58] Gonzalo, J. (1994): *Five Alternative Methods of Estimating long-run Equilibrium Relationships*. Journal of Econometrics, 60, 203–233.

[59] Engle, R. F. and Granger C. W. J. (1987): *Co-Integration and Error Correction: Representation, Estimation and Testing*. Econometrica, 55, 251–276.

[60] Engle, R. F. and Granger C. W. J.(1991): *Long-Run Economic Relationships: Readings in Cointegration*. Oxford University Press, New York.

[61] Granger, C. W. J. (1969): *Investigating Causal Relations by Econometric Models and Cross Spectral Methods*. Econometrica, 37, 424–438.

[62] Granger, C. W. J. (1981): *Some Properties of Time Series Data and Their Use in Econometric Model Specification*. Journal of Econometrics, 16, 121–130.

[63] Granger, C. W. J. (1986): *Developments in the Study of Cointegrated Economic Variables*. Oxford Bulletin of Economics and Statistics 48, 213–228.

[64] Granger, C. W. J. and Newbold, P. E. (1974): *Spurious Regression in Econometrics*. Journal of Econometrics, 2, 111–120.

[65] Kaldor, N. (1982): *The Scourge of Monetarism*. Oxford University Press, London.

[66] Kaldor, N. (1985): *How Monetarism Failed*. Challenge, 28(2), 4–13.

[67] Haghighat, J. (2011): *Endogenous and Exogenous Money: An Empirical Investigation from Iran*. Journal of Accounting, Finance and Economics, 1(1), 61–76.

[68] Hakkio, C. S. and Rush M. (1991): *Cointegration: How Short is the Long Run?*. Journal of International Money and Finance, 10(4), 571–581.

[69] Haldrup, N. (1994): *Heteroskedasticity in Nonstationary Time Series, Some Monte Carlo Evidence*. Statistical Papers, 35, 287–307.

[70] Hamao, Y. (1988): *An Empirical Investigation of the Arbitrage Pricing Theory*. Japan and the World Economy, 1, 45–61.

[71] Hamilton, J.D. (1983): *Oil and the Macroeconomy since World War II*. The Journal of Political Economy, 91(2), 228–248.

[72] Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press, Princeton.

[73] Hatanaka, M. (1996): *Time Series Based Econometrics: Unit Roots and Cointegration*. Oxford University Press, New York.

[74] Hendry, D. F. (1995): *Dynamic Econometrics*. Oxford University Press, Oxford.

[75] Hendry, D. F. and Juselius, K. (2001): *Explaining Cointegration Analysis: Part II*. The Energy Journal, 22(1), 75–120.

[76] Hondroyiannis, G. and Papapetrou, E. (2001): *Macroeconomic Influences on the Stock Market*. Journal of Economics and Finance 25(1), 33–49.

[77] Hylleberg, S., Engle, R. F., Granger, C. W. J. and Yoo, B. S. (1990): *Seasonal Integration and Cointegration*. Journal of Econometrics, 44(1-2), 215–238.

[78] Ibrahim, M. H. (1999): *Macroeconomic Variables and Stock Prices in Malaysia: an Empirical Analysis*. Asian Economic Journal, 13, 219–231.

[79] Jaffe, J. and Mandelkar, G. (1976): *The Fisher Effect for Risky Assets: An Empirical Investigation*. Journal of Finance 31, 447–456.

[80] Jarque, C. M. and Bera, A. K. (1987): *A Test for Normality of Observations and Regression Residuals*. International Statistical Review, 55, 163–172.

[81] Jefferis, K. and Okeahalam, C. (2000): *An Event Study of the Botswana, Zimbabwe and Johannesburg Stock Exchanges*. South African Journal of Business Management, 30(4), 131–140.

[82] Johansen, S. (1991): *Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models*. Econometrica, 59, 1551–1580.

[83] Johansen, S. (1996): *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.

[84] Johansen, S. (1988): *Statistical Analysis of Cointegration Vectors*. Journal of Economic Dynamics and Control, 12, 231–254.

[85] Johansen, S. (2000): *Modelling of Cointegration in the Vector Autoregressive Model*. Economic Modelling, 17, 359–373.

[86] Johansen, S. and Juselius, K. (1990): *Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money*. Oxford Bulletin of Economics and Statistics, 52, 169–210.

[87] Kim, K. and Schmidt, P. (1993): *Unit Root Tests with Conditional Heteroskedasticity*. Journal of Econometrics, 59, 287-300.

[88] Kim, T. H., Leybourne, S. J. and Newbold, P. (2002): *Unit Root Tests with a Break in Innovation Variance*, Journal of Econometrics, 109, 365–387.

[89] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1991): *Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root*. Journal of Econometrics, 54, 159–178.

[90] Kwon, C. S. and Shin, T. S. (1999): *Cointegration and Causality between Macroeconomic Variables and Stock Market Returns*. Global Finance Journal, 10(1), 71–81.

[91] Lavoie, M. (2005): *Monetary Base Endogeneity and the New Procedures of the Asset-Based Canadian and American Monetary Systems*. Journal of Post Keynesian Economics, 27(4), 689–708.

[92] Lee, B. S. (1992): *Causal Relations among Stock Returns, Interest Rates, Real Activity and Inflation*. Journal of Finance, 47, 1591–1603

[93] Lee, T. H. and Tse, Y. (1996): *Cointegration Tests with Conditional Heteroscedasticity*. Journal of Econometrics, 73, 401–410.

[94] Lekobane-Sikalao, O. L. and Lekobane, K. R. (2012): *Do Macroeconomic Variables Influence Domestic Stock Price Behavior in Emerging Markets? A Johansen Cointegration Approach to the Botswana Stock Market*. Journal of Economic and Behavioral Studies, 6(5), 363–372.

[95] Ljung, G. M. and Box, G. E. P. (1978): *On a Measure of a Lack of Fit in Time Series Models*. Biometrika, 65(2), 297–303

[96] Lütkepohl, H. (2004): *Recent Advances in Cointegration Analysis*. Economics Working Papers ECO2004/12, European University Institute.

[97] Lütkepohl, H. (2006): *New Introduction to Multiple Time Series Analysis*. Springer, New York.

[98] Lütkepohl, H. (2011): *Vector Autoregressive Models*. EUI Working Papers 30.

[99] Lütkepohl, H. and Krätzig, M. (2004): *Applied Time Series Econometrics*. Cambridge University Press.

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

[101] McMillan, D. G. (2001): *Cointegration Relationships between Stock Market Indices and Economic Activity: Evidence from U.S. Data*. Centre for Research into Industry, Enterprise, Finance and the Firm, CRIEFF Discussion Papers, 0104.

[102] Mahedi M. (2012): *Impact of the Macroeconomic Variables on the Stock Market Returns: The Case of Germany and the United Kingdom*. Global Journal of Management and Business Research 12(16), 23–34.

[103] Maysami, R. C. and Koh, T. S. (2000): *A Vector Error Correction Model of the Singapore Stock Market*. International Review of Economics and Finance, 9, 79–96.

[104] Maysami, R. C. and Sim, H. H. (2001a): *An Empirical Investigation of the Dynamic Relations between Macroeconomics Variable and the Stock Markets of Malaysia and Thailand*. Jurnal Pengurusan 20, 1–20.

[105] Maysami, R. C. and Sim, H. H. (2001b): *Macroeconomic Forces and Stock Returns: a General-to-Specific ECM Analysis of the Japanese and South Korean Markets*. International Quarterly Journal of Finance 1(1), 83–99.

[106] Maysami, R. C. and Sim, H. H. (2002): *Macroeconomics Variables and their Relationship with Stock Returns: Error Correction Evidence from Hong Kong and Singapore*. The Asian Economic Review 44(1), 69–85.

[107] Maysami, R. C., Howe, L. C. and Hamaz, M. A. (2004): *Relationship between Macroeconomic Variables and Stock Market Indices: Cointegration Evidence from Stock Exchange of Singapore's All-S Sector Indices*. Journal Penguruson, 24, 47–77.

[108] Moore, B. J. (1979): *The Endogenous Money Stock*. Journal of Post Keynesian Economics, 2(1), 49–70.

[109] Moore, B. J. (1983): *Unpacking the Post Keynesian Black Box: Bank Lending and the Money Supply*. Journal of Post Keynesian Economics, 5(4), 537–556.

[110] Moore, B. J. (1986): *How Credit Drives the Money Supply: The Significance of Institutional Developments*. Journal of Economic Issues, 20(2), 443–452.

[111] Moore, B. J. (1988): *The Endogenous Money Supply*. Journal of Post Keynesian Economics, 10(3), 372–385.

[112] Moore, B. J. (1989): *On the Endogeneity of Money Once More*. Journal of Post Keynesian Economics, 11(3), 479–487.

[113] Mukherjee, T. K. and Naka, A. (1995): *Dynamic Relations between Macroeconomic Variables and the Japanese Stock Market: an Application of a Vector Error Correction Model*. The Journal of Financial Research, 18(2), 223–237

[114] Muradoglu, G., Metin, K. and Argae, R. (2001): *Is There a Long-run Relationship Between Stock Returns and Monetary Variables: Evidence from an Emerging Market*. Applied Financial Economics, 7(6), 641–649.

[115] Naik, P. K. and Phadi, P. (2012): *The Impact of Macroeconomic Fundamentals on Stock Prices Revisited: Evidence from Indian Data*. Eurasian Journal of Business and Economics, 5(10), 25–44.

[116] Nell, K. S. (2000): *The Endogenous/Exogenous Nature of South Africa's Money Supply under Direct and Indirect Monetary Control Measures*. Journal of Post Keynesian Economics, 23(2), 313-329.

[117] Nelson, C. R. (1972): *The Prediction Performance of the FRB-MIT-PENN Model of the US Economy*. American Economic Review, 5, 902–917.

[118] Nelson, C. R. (1976): *Inflation and Rates of Return on Common Stocks*. Journal of Finance 31(2), 471–483.

[119] Nelson, C. R. and Plosser, C. I. (1982): *Trends and Random Walks in Macroeconomic Time Series*. Journal of Monetary Economics, 10, 139–162.

[120] Nelson, C. R. and Schwert, G. W. (1977): *On Testing the Hypothesis that the Real Rate of Interest is Constant*. American Economic Review, 478–486.

[121] Nasseh, A. and Strauss, J. (2000): *Stock Prices and Domestic and International Macroeconomic Activity: A Cointegration Approach*. Quarterly Review of Economics and Finance, 40(2), 229–245.

[122] Osterwald-Lenum, M. (1992): *A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Statistics*. Oxford Bulletin of Economics and Statistics, 54, 461–472.

[123] Onyemachi Maxwell Ogbulu, O. M., Abaenewe, Z. C. and Nnamocha, P. N. (2014): *Stock Market Price Behavior and Macroeconomic Variables in Nigeria: An Error Correction Analysis*. International Journal of Financial Economics, 2(2), 58–72.

[124] Padhan, P. C. (2007): *The Nexus between Stock Market and Economic Activity: an Empirical Analysis for India*. International Journal of Social Economics, 34, 741–753.

[125] Palley, T. I. (1987): *Bank Lending, Discount Window Borrowing, and the Endogenous Money Supply*. Journal of Post Keynesian Economics, 10, 282–301.

[126] Palley, T. I. (1994): *Competing Views of the Money Supply: Theory and Evidence*. Metroeconomica, 45, 67–88.

[127] Panagopoulos, Y. and Spiliotis, A. (1998): *The Determinants of Commercial Banks' Lending Behavior: Some Evidence for Greece*. Journal of Post Keynesian Economics, 20(4), 649–672.

[128] Panetta, F. (2002): *The Stability of the Relation between the Stock Market and Macroeconomic Forces*. Economic Notes, 31, 417–450.

[129] Pantula, S. G. (1988): *Estimation of Autoregressive Models with ARCH Errors*. Sankhya B50, 119–138.

[130] Pearce, D. K. (1983): *Stock Prices and the Economy*. Federal Reserve Bank of Kansas City Economic Review, 722.

[131] Perron, P. (1989): *The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis*. Econometrica, 57(6), 1361–1401.

[132] Perron, P. (1990): *Testing for a Unit Root in a Time Series with a Changing Mean*. Journal of Business and Economic Statistics, 8(2), 153–162.

[133] Pfaff, B. (2006): *Analysis of Integrated and Cointegrated Time Series*. Springer, New York.

[134] Phillips, P. C. B. (1986): *Understanding Spurious Regression in Econometrics.* Journal of Econometrics, 33, 311–340.

[135] Phillips, P. C. B.(1987): *Time Series Regression with a Unit Root*. Econometrica, 55, 277–301.

[136] Phillips, P. C. B.(1988): *New Tools for Understanding Spurious Regressions*. Econometrica, 66, 1299–1325.

[137] Phillips, P. C. B. (1991): *Optimal Inference in Cointegrated Systems*. Econometrica, 59, 283–306.

[138] Phillips, P. C. B. and Ouliaris, S. (1990): *Asymptotic Properties of Residual Based Tests for Cointegration*. Econometrica, 58, 73–93.

[139] Phillips, P. C. B. and Perron, P. (1988): *Testing for a Unit Root in Time Series Regression*. Biomètrica, 75, 335–46.

[140] Phillips, P. C. B. and Xiao, Z. (1998): *A Primer on Unit Root Testing*. Journal of Economic Surveys, 12(5).

[141] Pollin, R. (1991): *Two Theories of Money Supply Endogeneity: Some Empirical Evidence*. Journal of Post Keynesian Economics, 13(3), 366–396.

[142] Poon, S. H. and Taylor, S. J. (1991): *Macroeconomic Factors and the UK Stock Market*. Journal of Business and Accounting, 18, 619–636.

[143] Rahman, A. A., Noor, Z. M. S. and Fauziah, H. T. (2009): *Macroeconomic Determinants of Malaysian Stock Market*. African Journal of Business Management, 3(3), 95–106.

[144] Rao, B. B. (1994): *Cointegration for the Applied Economist*. St. Martin's Press, New York.

[145] Ratanapakorn, O. and Sharma, S. C. (2007): *Dynamic Analysis between the US Stock Return and Macroeconomic Variables*. Applied Financial Economics, 17(4), 369–377.

[146] Richards, A. J. (1996): *Volatility and Predictability in National Stock Markets: How do Emerging Markets Differ?* Internationally Monetary Fund Staff papers, 42, 461–501.

[147] Robert, D. G. (2008): *Effect of Macroeconomic Variables on Stock Market Returns for four Emerging Economies: Brazil, Russia, India and China*. International Business and Economics Research Journal, 7(3), 1–8.

[148] Ross, S. A. (1976): *The Arbitrage Theory of Capital Asset Pricing*. Journal of Economic Theory 13, 341–360.

[149] Runkle, D. E. (1987): *Vector Autoregressions and Reality*. Journal of Business and Economic Statistics, 5(4), 437–442.

[150] Schwarz, G. (1978): *Estimating the Dimension of a Model*. Annals of Statistics, 6, 461–4

[151] Schwert, W. G. (1989): *Why does Stock Market Volatility Change over Time?* Journal of Finance, 44, 1115–1153.

[152] Shiller, R. J. and Beltratti, A. E. (1992). *Stock Prices and Bond Yields: Can Their Comovements Be Explained in Terms of Present Value Models?*. Journal of Monetary Economics 30, 25–46.

[153] Sims, C. A. (1972): *Money, Income, and Causality*. American Economic Review, 62(4), 540–52.

[154] Sims, C. A. (1980): *Macroeconomics and Reality*. Econometrica, 48, 1–48.

[155] Sims, C. A., Stock, J. H. and Watson M. W. (1990): *Inference in Linear Time Series Models with Some Unit Roots*. Econometrica, 58, 113–144.

[156] Stock, J. H. (1987): *Asymptotic Properties of Least Squares Estimation of Cointegrating Vectors*. Econometrica, 55, 1035–1056.

[157] Stock, J.H. (1994): *Units Roots, Structural Breaks and Trends*. in R.F. Engle and D.L. McFadden (eds.), Handbook of Econometrics, Volume IV. North Holland, New York.

[158] Stock, J. H. and Watson, M. W. (1988): *Testing for Common Trends*. Journal of the American Statistical Association, 83, 1097–1107

[159] Stock, J. H. and Watson, M. W. (2001): *Vector Autoregressions*. Journal of Economic Perspectives, 15, 101–115.

[160] Tsay, R. S. (2001): *Analysis of Financial Time Series*. John Wiley & Sons, New York.

[161] Tsay, R. S. (2014): *Multivariate Time Series Analysis: with R and Financial Applications*. John Wiley & Sons, New York.

[162] Tsoukalas, D. (2003): *Macroeconomic Factors and Stock Prices in the Emerging Cypriot Equity Market*. Managerial Finance, 29, 87–92.

[163] Vera, A. P. (2001): *The Endogenous Money Hypothesis: Some Evidence From Spain (1987-1998)*. Journal of Post Keynesian Economics, 23(3), 509–526.

[164] Wallis, K. F. (1974): *Seasonal Adjustment and Relations between Variables*. Journal of the American Statistical Association, 69, 18–31.

[165] Wasserfallen, W. (1989): *Macroeconomic News and the Stock Market*. Journal of Banking and Finance, 13, 613–626.

[166] Watson, M. W. (1994): *Vector Autoregressions and Cointegration*. in Handbook of Econometrics, Volume IV. R.F. Engle and D. McFadden (eds.). Elsevier Science Ltd., Amsterdam.

[167] Wooldridge, J. M. (2003): *Introductory Econometrics: A Modern Approach*. 5th edn., Cengage Learning, Ohio.

[168] Zivot, E. and Wang, J. (2005): *Modelling Financial Time Series with S-PLUS*. 2nd edn., Springer, New York.

[169] Zivot, E. and Andrews, D. W. K. (1992): *Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis*. Journal of Business and Economic Statistics, 10(3), 251–270.