Contents

List of Tables

List of Figures

Acknowledgement

First and foremost, I thank the Almighty God for His abundant mercies, blessings, protection and how far He has brought me and for granting me the opportunity and strength to undertake this wonderful programme successfully.

I feel privileged to have worked with my supervisor, professor Umberto Triacca. I thank him for his guidance, patience and supervision throughout this research work. I have really learned a lot from him. I will also like to thank the MathMod team especially the coordinator, professor Rubino Bruno and Doctore Danilo Larivera for their advice and assistance to my welfare and stay in L'Aquila. I appreciate the effort of Hamburg coordinating team, Professor Dr. Ingenuin Gasser, Ms. Astriz Benz and Dr. Felaco for their warm reception and guidance I received for my stay in Hamburg University, Germany. Sincere thanks to the many lecturers who taught me and my colleagues in the MathMod programme whose stay together as a family made home out from home.

Finally, I am indebted to my family for their invaluable support and prayers for me especially my father, mother, sister and brother back home in Ghana and all other aspect to aid me achieve my aims and also my friends.

Abstract

The relationship between the exchange rate and economic development is certainly an important subject, from both a positive (descriptive) and a normative (policy prescription) perspective. Several developing countries that have implicitly or explicitly fixed their exchange rates to the currency of another country (say, the U.S. dollar) and whose inflation rates are higher than that of the foreign country (the United States) often experience persistent current account deficits and eventual devaluations of their currencies. Other developing countries grow exceptionally fast and often face the opposite pressure on their currencies. A high economic growth rate is most likely accompanied by a high investment rate, and high export growth as well. Successful exports produce current account surpluses, resulting in nominal appreciation pressure on the currency unless the central bank intervenes in the foreign exchange market and accumulates foreign reserves. This thesis tries to investigate the possibility of Granger causality between the logarithms of GDP, exports and exchange rate in twenty - seven (27) African developing countries from 1965 - 2010. We examine the panel causality between all the 27 countries at once, and then performed the individual panel analysis for each country using Seemingly Unrelated Regression estimation.

Chapter 1

Introduction

1.1 Background of Study

The impart of export and exchange rate on economic growth (GDP) is a very essential issue to policy makers of every country, be it, developed, developing or under - developed countries. Many developing countries especially those in Africa, have inherently or dubiously fixed their exchange rate to that of another country's currency, say US $ in order to reduce their exchange rate. On the other hand, some delveloping countries also experience exceptionally fast growth which leads to pressure on their currencies. A high economic growth rate is most likely followed by a high investment rate, which in turn leads to high export growth as well. Successful exports ensures a current account surpluses, resulting in nominal appreciation pressure on the currency unless the central bank intervenes in the foreign exchange market and accumulates foreign reserves. Even if the intervention maintains the fixed exchange rate, unsterilized intervention results in inflation, and the real exchange rate appreciates anyway. That is, successful economic development results in a currency appreciation with improvement in the standard of living, while failure in economic development often results in a sharp currency depreciation

The literature on exports and economic growth has its source in the late 1970s. The methodology of the early studies relies on correlation coefficients between export growth and economic growth (as in Michaely (1977 [34]) Michalopoulos and Jay (1973) [35]). In the 1980s, most studies used the Granger causality test method to investigate lead - and-lag relations. Notable examples include Chow (1987) [9] and Jung and Marshall (1985) [28]. In the 1990s, the development of the concepts of unit root and cointegration added twist to studies employing the causality test (see for example, Bahmani- Oskooee et al. (1991) [1], Sharma et al. (1991) [45], BahmaniOskooee and Alse (1993) [2], Sharma and Dhakal (1994) [44], Ghartey (1993) [20], Xu (1996) [52], Riezman et al. (1996) [40], Huang, Oh and Yang (2000) [26], and Shan and Sun (1998) [46]). And finally, in the

2000's, where Kónya (2006) [31] proposed bootstrap panel Granger causality, which does not requir testing the variables for unit root and cointegration, in this case the variables being used in their levels. Secondly, the tool considers the existence of contemporaneous correlations across the variables and offers additional panel information (the equations composes a SUR system).

Loosely speaking, export growth can promote economic growth and vice versa. Thetheoretical justification for these hypotheses is discussed as follows.
From the growth - theory literature point of view, export expansion is the key factor promoting economic growth. There are various explanations that have been put forward to relate these two variables to each other. First, the growth of exports has a stimulating effect on total factor productivity growth through its positive impact on higher rates of capital formation [10]. Second, the growth of exports helps relax the foreign exchange constraints, thereby facilitating imports of capital goods and hence faster growth. Third, competition from overseas ensures an efficient price mechanism that fosters optimum resource allocation and increases the pressure on industries that export goods to keep costs relatively low and to improve technological change, thereby promoting economic growth. Clearly, these arguments lead to hypothesize that exports contribute positively to economic progress.

In contrast to the export - led growth hypothesis, it can also be argued that causality runs from the growth of output to the growth of exports. When we consider a growing economy, some industries face substantial changes in terms of learning and technological innovation, which are related to the accumulation of human capital, manufacturing experiences and the technology transfer or real capital accumulation arising from foreign direct investment. Such unbalanced growth has nothing to do with outward - oriented policies, i.e., output will still continue to grow even in the absence of these policies. Under such unbalanced growth, the growth of domestic demand will lag behind the growth of output in these prosperous industries and it is likely that the producers will sell their goods in overseas markets. Therefore, economic growth will promote the growth of exports.
Another plausible hypothesis is that negative causality runs from output growth to export growth. This would be likely to occur if consumer demand were concentrated in exportable and non-traded goods in which case an increase in domestic demand would induce an increase in output but a decrease in exports. As a result, output growth will lead to a reduction in the growth of exports. If an increase in exports arises as a result of inward foreign direct investment, the growth of exports will reduce the growth of output due to various distortions (Bhagwati (1979)[3]), and it is therefore easier to identify the negative relationship between the growth of output and the growth of exports.

1.2 Literature Review

There is a vast amount of empirical literature on this issue. The most recent and most comprehensive survey of this literature is due to Giles and Williams (2000) [25] who reviewed more than one hundred and fifty export - growth applied papers published between 1963 and 1999. These papers fall into three groups. The first group of studies is based on cross - country rank correlation coefficients, the second applies cross-sectional regression analysis, and the third uses time series techniques on a country by-country basis. Two thirds of the papers belong to this third group, and more than seventy of these are based on the concept of Granger causality and on various tests for it.

"Disagreements persist in the empirical literature regarding the causal direction of the effects of trade openness on economic growth. Michaely (1977) [34], Feder (1982) [18], Marin (1992) [36], Thornton (1996) [50] found that countries exporting a large share of their output seem to grow faster than others. The growth of exports has a stimulating influence across the economy as a whole in the form of technological spillovers and other externalities. Models by Grossman and Helpman (1991) [21], Rivera-Batiz and Romer (1991) [41], Romer (1990) [42] posit that expanded international trade increases the number of specialized inputs, increasing growth rates as economies become open to international trade. Buffie(1992) [4] considers how export shocks can produce export-led growth" (Ribeiro Ramos, 2001) [43]. "Oxley (1993) [37], using Portuguese data, finds no support for the ELG hypothesis, quite the reverse, adding fuel to the controversy concerning programmes for growth. Export growth is often considered to be a main determinant of the production and employment growth of an economy. This so - called hypothesis of export-led growth (ELG) is, as a rule, substantiated by the following four arguments" (Balassa [5], 1978; Bhagwati, 1978 [6]; Edwards, 1998 [15]). "First, export growth leads, by the foreign trade multiplier, to an expansion of production and employment. Second, the foreign exchange made available by export growth allows the importation of capital goods which, in turn, increase the production potential of an economy. Third, the volume of and the competition in exports markets cause economies of scale and an acceleration of technical progress in production. Fourth, given the theoretical arguments mentioned above, the observed strong correlation of export and production growth is interpreted as empirical evidence in favor of the ELG hypothesis" (Ribeiro Ramos, 2001 [43]). "Export expansion and openness to foreign markets is viewed as a key determinant of economic growth because of the positive externalities it provides. For example, firms in a thriving export sector can enjoy the following benefits: efficient resource allocation, greater capacity utilization, exploitation of economies of scale, and increased technological innovation stimulated by foreign market competition" (Helpman and krugman, 1985 [27]). "In the GLE case, export expansion could be stimulated by productivity gains caused by increase in domestic levels of skilled-labor and technology (Bhangwati, 1988 [6]; Krugman, 1984 [30]). Neoclassical trade theory typically stresses the causality that runs from home-factor endowments and productivity to the supply of exports (Findlay, 1984 [19]). The product life cycle hypothesis developed by Vernon (1996) [51] has also attracted considerable attention among international trade theorists in recent years. Segerstrom et al. (1990) [47], for example, use the product life cycle hypothesis as a basis for analyzing north - south trade in which research and development competition between firms determines the rate of product innovation in the north" (Ribeiro Ramos, 2001)[43]..

A more recent work is by László Kónya [31], his paper investigates the possibility of Granger causality between the logarithms of real exports and real GDP in twenty-four OECD countries from 1960 to 1997. A new panel data approach is applied which is based on SUR systems and Wald tests with country specific bootstrap critical values. Two different models are used. A bivariate (GDP-exports) model and a trivariate (GDP-exports-openness) model, both without and with a linear time trend. In each case the analysis focusses on direct, one-period-ahead causality between exports and GDP. The results indicate one-way causality from exports to GDP in Belgium, Denmark, Iceland, Ireland, Italy, New Zealand, Spain and Sweden, one-way causality from GDP to exports in Austria, France, Greece, Japan, Mexico, Norway and Portugal, two-way causality between exports and growth in Canada, Finland and the Netherlands, while in the case of Australia, Korea, Luxembourg, Switzerland, the UK and the USA there is no evidence of causality in either direction.

1.3 Motivation

Although previous empirical work has been concentrated on a large number of both developed and developing countries using export and GDP or exchange rate and GDP. The literature on this subject has largely neglected these three economic variables all together owing to the non - availability of consistent data or lack of interest. In this study, we try ascertain the short run dynamics and long term effect on the export, exchange rate and the GDP in developing countries (Africa).

Our main focus, is to perform a Granger causality analysis for the three economic varaibles, Export, exchange rate and GDP, and also to build Vector Error Correction Model (VECM) which will then enable us to study the short - run dynamics and the long - run causal relationships between our three economic variables. In particular, using the vector error correction model, we seek to address the following research questions.

- What are the short and long run relationship between export, exchange rate and GDP for the overall panel data?

- What are the short and long run relationship between export, exchange rate and

GDP for each country?

- We will do the following for the overall panel data and for each country:
  - Does export and exchange rate jointly Granger cause GDP?
  - Does export and GDP jointly Granger cause exchange rate?
  - Does GDP and exchange rate jointly Granger cause export?
  - Does export Granger cause exchange rate and GDP?
  - Does exchange rate Granger cause export and GDP?
  - Does GDP Granger cause export and exchange rate?

Chapter 2

Granger Causality Concept

The identification of causal relationships is an important part of scientific research and essential for understanding the consequences when moving from empirical findings to actions. At the same time, the notion of causality has shown to be evasive when trying to formalize it. Among the many properties a general definition of causality should or should not have, there are two important aspects that are of practical relevance:
Temporal precedence: causes precede their effects;
Physical influence: The cause makes unique changes in the effect. In other words, the causal series contains unique information about the effect series that is not available otherwise.
In time series analysis, most approaches to causal inference make use of the first aspect of temporal precedence. One the one hand, controlled experiments are often not feasible in many time series applications and researches may be reluctant to think in these terms. On the other hand, temporal precedence is readily available in time series data.

2.1 Digress

2.1.1 Stochastic Process

Many financial and economic variables in traditional mathematical finance models are based on stochastic processes. These variables evolve over time and behave randomly fully or partly with some known distributional properties. In general, a real valued collection of variables $\{x(t) : t \in \mathcal{T}, \mathcal{T} \subseteq R\}$ that move discretely or continuously in time t and unpredictably or at least partly random are said stochastic process. Each run of a stochastic process gives a realization of the process. Our time series data $\{x(t) : t \in \mathcal{T}\}$ is a finite part of a realization of a stochastic process.

2.1.1.1 Stationary Stochastic Process

The mathematical theory of stochastic processes try to determine the classes of processes for which a unified theory can be developed. One of such important classes is stationary process. Time series data may exhibit stationarity or not at their levels. For the estimation of many models, it becomes imperative to detrend the data if they exist before certain statistical inference can be made.

Definition 1: Given a probability space $(\Omega, \mathcal{F}, P)$, a discrete stochastic process is a sequence of random variables indexed by t , $t \in \mathbb{Z}$, i.e, $\{x_t : t \in \mathbb{Z}\}$ all defined on $(\Omega, \mathcal{F}, P)$. A time series $\{x_t : t \in \mathbb{Z}\}$ can be considered as a realization of a final part of a stochastic process $\{x_t : t \in \mathcal{T}\}$ with $\mathcal{T} \subset \mathbb{Z}$.

Definition 2: A stochastic process $\{x_t : t \in \mathcal{T}, \mathcal{T} \subset \mathbb{Z}\}$ is said to be stationary if:

- $E(x_t)^2 < \infty \qquad \forall t \in \mathbb{Z}$
- $E(x_t) = \mu \qquad \forall t \in \mathbb{Z}$
- $Cov(x_{t_1}, x_{t_2}) = Cov(x_{t_{1+h}}, x_{t_{2+h}}) \qquad \forall t_1, t_2 \in \mathbb{Z}$

As used in many literature, this definition can be termed as covariance stationary, weak stationarity or stationarity in the wider sense or second order stationarity.

In other words, the process $\{x_t : t \in \mathbb{Z}\}$ is stationary if it its mean and variance are constant over time and the covariance depends only on the lag distance between the two time periods but not the actual time t itself. Intuitively, stationarity of $\{x_t : t \in \mathbb{Z}\}$ means that a certain type of statistical equilibrium is achieved and the distribution of $\{x_t : t \in \mathbb{Z}\}$ does not change much.

Definition 3: A stationary process $\{x_t : t \in \mathcal{T}, \mathcal{T} \subset \mathbb{Z}\}$ is called strictly stationary if the joint distribution of the random vectors $(x_{t_1}, ..., x_{t_n})'$ and $(x_{t_{1+h}}, ..., x_{t_{n+h}})'$ are the same for all positive integers n and for all indices $\{t_1, ..., t_n\}, h, t \in \mathbb{Z}$. Strict sationarity intuitively means that the graphs over two equal-length time intervals of a realization of the time series exhibit similar statistical characteristics.

Definition 4: A stochastic process $\{x_t : t \in \mathcal{T}, \mathcal{T} \subset \mathbb{Z}\}$ is said to be Gaussian Stochastic process if and only if the joint distribution functions of $x_{t_1}, x_{t_2}, ..., x_{t_n}$ for all finite subsets $(t_1, ..., t_n)$ is normal. Generally, strong stationarity does not imply weak stationarity and vice versa. However, if the process $\{x_t : t \in \mathcal{T}, \mathcal{T} \subset \mathbb{Z}\}$ is a Gaussian process, then strong and weak stationarity are equivalent. Suppose $\{x_t : t \in \mathcal{T}, \mathcal{T} \subset \mathbb{Z}\}$ is strictly stationary, then it is also covariance stationary if $E(x_t)^2 < \infty$.

2.2 Granger Causality

In the following, we consider two weakly stationary stochastic processes, $\{x_t : t \in \mathcal{T}, \mathcal{T} \subset \mathbb{Z}\}$ and $\{y_t : t \in \mathcal{T}, \mathcal{T} \subset \mathbb{Z}\}$. Let $\mathcal{F}_t$ be the total information set available at time t. This information set includes, the set of all current and past values of x, i.e. $\bar{x_t} := \{x_t, x_{t-1}, ..., x_{t-k}, ..\}$ and analogously of y, i.e. $\bar{y_t} := \{y_t, y_{t-1}, ..., y_{t-k}, ..\}$. Let $\sigma(\cdot)$ be the variance of the corresponding forecast error. For such a situation, C.W.J. GRANGER (1969) proposed the following definition of causality between x and y:

Definition 5: x is said *not* to Granger-cause y if for all h > 0

$$\mathcal{F}(y_{t+h}|\Omega_t) = \mathcal{F}(y_{t+h}|\Omega_t - \bar{x_t})$$

where $\mathcal{F}$ denotes the conditional distribution and $\Omega_t - \bar{x_t}$ is all the information in the universe except series x. In plain words, x is said to not Granger-cause y if x_t cannot help predict future y_t.

Remarks:
The whole F is generally difficult to handle empirically and we turn to conditional expectation and variance.

Definition 6: A redefined defintion become as below:
x is said *not* to Granger-cause y if for all h > 0

$$E(y_{t+h}|\Omega_t) = E(y_{t+h}|\Omega_t - \bar{x_t})$$

Remarks:
• It is defined for all $h > 0$ and for for only $h = 1$. Causality at different h does not imply each other.They are neither sufficient nor necessary
• Ω_t contains all the information in the universe up to time t that excludes the potential ignored common factors problems. the question is: how to measure Ω_t in practice? The unobserved common factors are always a potential problrm for any finite information set.

Definition 7: A redefined defintion become as below:
x does not Granger cause y with respect to information J_t, if

$$E(y_{t+1}|J_t) = E(y_{t+1}|J_t - \bar{x_t})$$

Where $\Omega_t \supset J_t \supset \bar{x}_t \cup \bar{y}_t$

Definition 8: Let $P(y_{t+1}|J_t)$ be the optimal forecast of the y_{t+1} based on the information set F_t and $P(y_{t+1}|J_t - \bar{x}_t)$ be the forecast ofthe y_{t+1} based on the information set $J_t - \bar{x}_t$. We say that x is not (simply) **Granger causal** to y if and only if

$$P(y_{t+1}|J_t) = P(y_{t+1}|J_t - \bar{x}_t)$$

this is equivalent to saying that, the application of an optimal prediction function leads to

$$\sigma^2(y_{t+1}|J_t) = \sigma^2(y_{t+1}|J_t - \bar{x}_t) = E(y_{t+1} - P(y_{t+1}|J_t - \bar{x}_t))^2$$

i.e. if future value of y can be predicted better, i.e. with a smaller forecast error variance, if current and past values of x are used.

Definition 9: We say that x is not **instantaneously Granger causal** to y if and only if

$$P(y_{t+1}|\{J_t, x_{t+1}\}) = P(y_{t+1}|J_t)$$

that is, the application of an optimal linear prediction function leads to

$$\sigma^2(y_{t+1}|\{J_t, x_{t+1}\}) < \sigma^2(y_{t+1}|J_t),$$

i.e. if the future value of y, y_{t+1}, can be predicted better, i.e. with a smaller forecast error variance, if the future value of x,x_{t+1}, is used in addition to the current and past values of x.

Definition 10: There is **feedback** between x and y if x is causal to y and y is causal to x.

Feedback is only defined for the case of simple causal relations. The reason is that the direction of instantaneously causal relations cannot be identified without additional information or assumptions. Thus, the following theorem holds:

Theorem:

x is instantaneously causal to y if and only if y is instantaneously causal to x.
According to this definition there are eight different, exclusive possibilities of causal relations between two stochastic process:

- x and y are independent: (x, y)
- There is only instantaneous causality: $(x - y)$
- x is causal to y, without instantaneous causality: $(x \rightarrow y)$

- y is causal to x, without instantaneous causality: $(x \leftarrow y)$
- x is causal to y, with instantaneous causality: $(x \implies y)$
- y is causal to x, with instantaneous causality: $(x \Leftarrow y)$
- There is feedback without instantaneous causality: $(x \leftrightarrow y)$
- There is feedback with instantaneous causality: $(x \Leftrightarrow y)$

Remark:
Sometimes econometrians use the shorter terms "causes" as shorthand for "Granger causes". You should notice, however, that Granger causality is not causality in a deep sense of the word. It just talk about linear prediction, and it only has "teeth" if one thing happens before another. (In other words if we only find Granger causality in one direction). In economics you may often have that all variables in the economy reacts to some unmodeled factor (the Gulf war) and if the response of x_t and y_t is staggered in time you will see Granger causality even though the real causality is different. There is nothing we can do about that (unless you can experiment with the economy) - Granger causality measures whether one thing happens before another thing and helps predict it - and nothing else.

2.3 Granger Causality in Vector Autoregressive (VAR) Framework

Let $\{y_t = (y_{1t}, y_{2t}, ..., y_{Kt})' : t \in \mathbb{Z}\}$ be a K-variate stochastic process. Suppose that $\{y_t = (y_{1t}, y_{2t}, ..., y_{Kt})' : t \in \mathbb{Z}\}$ follows a Vector Autoregressive model of order p, VAR(p)

$$y_t = \nu + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + u_t, \qquad t = 0, \pm 1, \pm 2, ..., \quad or$$

$$y_t = \nu + A(B) y_{t-1} + u_t \qquad t = 0, \pm 1, \pm 2, ...,$$

where the $A_i = [a_{lj}^{(i)}]$ are fixed $(K \times K)$ cofficient matrices, $By_t = y_{t-1}$, $A(B) = A_1 B + A_2 B^2 + ... + A_p B^p$, $\nu = (\nu_1, ..., \nu_K)'$ is a fixed $(K \times 1)$ vector of intercept terms allowing for the possibility of a nonzero mean $E(y_t)$. Finally,$u_t = (u_{1t}, u_{2t}, ..., u_{Kt})'$ is a K-dimensional white noise process, that is, $E(u_t) = 0$, $E(u_t u_t') = \Sigma_u$ and $E(u_t u_s') = 0$ for $s \neq t$. The covariance matrix Σ_u is assumed to be nonsingular.

In a country- by - country analysis the possibility of Granger causality between our variables LEXP, LGDP and LEXC can be studied using the following trivariate finite -

order vector autoregressive (VAR) model:

$$
\begin{aligned}
y_{i,t} &= \alpha_{1,i} + \epsilon_{1,i,t} + \sum_{l=1}^{n_y} \beta_{1,i,l} y_{i,t-1} + \sum_{l=1}^{n_x} \gamma_{1,i,l} x_{i,t-1} \\
&+ \sum_{l=1}^{n_z} \eta_{1,i,l} z_{i,t-1} \\
x_{i,t} &= \alpha_{2,i} + \epsilon_{2,i,t} + \sum_{l=1}^{n_y} \beta_{2,i,l} y_{i,t-1} + \sum_{l=1}^{n_x} \gamma_{2,i,l} x_{i,t-1} \\
&+ \sum_{l=1}^{n_z} \eta_{1,i,l} z_{i,t-1} \\
z_{i,t} &= \alpha_{3,i} + \epsilon_{1,i,t} + \sum_{l=1}^{n_y} \beta_{3,i,l} y_{i,t-1} + \sum_{l=1}^{n_x} \gamma_{3,i,l} x_{i,t-1} \\
&+ \sum_{l=1}^{n_z} \eta_{3,i,l} z_{i,t-1}
\end{aligned}
\tag{2.1}
$$

where index i refers to the country $(i = 1, ..., N)$, t to the time period $(t = 1, ..., T)$, l to the lag and n_x, n_y, n_z refers respectively to the optimal lag for variables x, y and z. and $\epsilon_{1,i,t}, \epsilon_{2,i,t}, \epsilon_{3,i,t}$ are assumed to be white - noise errors that may be correlated for a given country, but not accross countries.

Consider is a K- *dimensional VAR(p)* process of the form

$$
y_t = \Lambda + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \epsilon_t \tag{2.2}
$$

where $y_t = (y_{1t}, ..., y_{kt})'$ is a $(K \times 1)$ column vector of endogenous variables, A_i are fixed $(K \times K)$ matrix of coefficient parameter which can be represented as

$$
A_i = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{K1} & a_{K2} & \cdots & a_{KK} \end{pmatrix}
$$

$\Lambda = (\Lambda_1, ..., \Lambda_K)'$ is a fixed $(K \times 1)$ vector of intercept terms indicating a possible non - zero mean with $E(y_t) \neq 0$ and finally $\epsilon_t = (\epsilon_{1t}, ..., \epsilon_{1K})$ is a K - *dimensional* white noise process with zero mean, $E(\epsilon_t) = 0$ and a non singular covariace matrix $\Sigma_u = E(\epsilon_t \epsilon_t'), \forall t$.

In order to determine the Granger-causal relationships between the variables of the K - dimensional VAR process y_t, suppose it has the canonical MA representation

$$
y_t = \mu + \Phi(B) u_t = \mu + \sum_{i=1}^{\infty} \Phi_i u_{t-i} \qquad\qquad \Phi_0 = I_K
$$

where $\Phi_i = [\phi_{k,j,i}]$. A necessary and sufficient condition for variable x not Granger-cause variable y is that $\Phi_{jk,i} = 0$ for $i = 1, 2, 3, ...$

If there are only two variable, or two-group of variables, x_t and y_t, then a necessary and sufficient condition for variable x not to Granger - cause variable y is that $A_{jk,i} = 0$ for $i = 1, 2, 3, ...$ The condition is good for all forecast horizon, h.
Note that for a VAR(1) process with dimension equal or greater than 3, $A_{jk,i} = 0$ for $i = 1, 2, 3, ...$ is sufficient for non-causality at $h = 1$ but insufficient for $h > 1$. Variable k might affect variable j in two or more period in the future via the effect through other variables.

Granger causality is particularly easy to deal with in VAR models. Assume that our data can be described by the model

$$\begin{bmatrix} y_{1t} \\ zy_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} + \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} \\ a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + ... + \begin{bmatrix} A_{11}^{(p)} & A_{12}^{(p)} & A_{13}^{(p)} \\ A_{21}^{(p)} & A_{22}^{(p)} & A_{23}^{(p)} \\ A_{31}^{(p)} & A_{32}^{(p)} & A_{33}^{(p)} \end{bmatrix} \begin{bmatrix} y_{1t-p} \\ y_{2t-p} \\ y_{3t-p} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{bmatrix}$$

Also assume that

$$\Sigma_u = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12}' & \sigma_{22} & \sigma_{23} \\ \sigma_{13}' & \sigma_{23}' & \sigma_{33} \end{bmatrix}$$

In this model it is clear (convince yourself!) that y_3 does not Granger cause y_1 with respect to the information set $J_t = \bar{Y_1} \cup \bar{Y_2} \cup \bar{Y_3}$, if $a_{23}^{(i)} = 0; i = 1, ..., p$. Note that this is the way you will test for Granger causality. Usually you will use the VAR approach if you have an econometric hypothesis of interest that states that x_t Granger causes y_t but y_t does not Granger cause x_t. *Sims (1972)* is a paper that became very famous because it showed that money Granger causes output, but output does not Granger cause money. (This was in the old old days when people still took monetarism seriously, and here was a test that could tell whether the Keynesians or the monetarists were right!!). Later *Sims* showed that this conclusion did not hold if interest rates were included in the system. This also shows the major drawback of the Granger causality test - namely the dependence on the right choice of the conditioning set. In reality one can never be sure that the conditioning set has been chosen large enough (and in short macro-economic series one is forced to choose a low dimension for the VAR model), but the test is still a useful (although not perfect) test.

2.3.1 Remark

To summarize,

- At this point,some words od caution seem appropriate. The term *"causality"* suggests a cause and effect relationship between two sets of variables. This interpretation is problematic with respect to instantaneous causality because this term only describ es a nonzero correlation b etween two sets of variables. It does not say anything about the cause and effect relation. The direction of instantaneous causation cannot be derived from the MA or VAR representation of the process but must be obtained from further knowledge on the relationship between the variables. Such knowledge may exist in the form of an economic theory.

- Although a direction of causation has been defined in relation with Granger-causality it is problematic to interpret the absence of causality from y_{2t} to y_{1t} in the sense that variations in y_{2t} will have no effect on y_{1t}. To see this consider, for instance, the stable bivariate VAR(1) system

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

In this system, y_2 does not Granger-cause y_1. However, the system may be multiplied by some nonsingular matrix $\begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix}$ so that we get the new transformation:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 0 & -\beta \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}$$

where $\gamma_{11} := \alpha_{11} + \alpha_{21}\beta$, $\gamma_{12} := \alpha_{22}\beta$, $\gamma_{21} := \alpha_{21}$, $\gamma_{22} := \alpha_{22}$ and $(v_{1t}, v_{2t})' := \beta(u_{1t}, u_{2t})'$.

- In other words, the stochastic interrelationships between the random variables of the system can either be characterized by any of the two equations above, although the two representations have quite different physical interpretations. Thus, the lack of a Granger-causal relationship from one group of variables to the remaining variables cannot necessarily be interpreted as lack of a cause and effect relationship. It must be remembered that a VAR or MA representation characterizes the joint distribution of sets of random variables. In order to derive cause and effect relationships from it, usually requires further assumptions regarding the relationship between the variables involved.

- Further problems related to the interpretation of Granger- causality result from restricting the information set to contain only past and present variables of the system rather all the information in the universe. Only if all other information inthe universe is irrelevant for the problem at hand, the reducttion of the information set is of no consequence.

I think that the Granger causality tests are most useful in situations where one is willing to consider 2-dimensional systems. If the data are reasonably well described by a 2-dimensional system ("no $z_t$variables")the Granger causality concept is most straightforward to think about and also to test. By the way, be aware that there are special problems with testing for Granger causality in co-integrated relations (see Toda and Phillips (1991))

In summary, Granger causality tests are a useful tool to have in a toolbox, but should be handled with care. It will very often be hard to find any clear conclusions unless the data can be described by a simple "2-dimensional" system (since the test may be between 2 vectors the system may not be 2-dimensional is the usual sense), and another potentially serious problem may be the choice of sampling period: a long sampling period may hide the causality whereas for example VAR-systems for monthly data may give you serious measurement errors (e.g. due to seasonal adjustment procedures).

2.4 Causality Test

The Granger causality test is a *statistical hypothesis test* for determining whether one Stochastic process is useful in forecasting another, first proposed in 1969 by C. W. J. Granger. Ordinarily, regressions reflect "mere" correlations, but Clive Granger argued that causality in economics could be tested for by measuring the ability to predict the future values of a time series using prior values of another time series. Since the question of "true causality" is deeply philosophical, and because of the *post hoc ergo propter hoc* fallacy of assuming that one thing preceding another can be used as a proof of causation, econometricians assert that the Granger test finds only "predictive causality".

In 1972, CHRISTOPHER A. SIMS was the first to propose a test for simple Granger causal relations. This test was based on the moving average representation. However, some problems occurred with this procedure. Therefore, it is hardly applied today and will not be discussed here. THOMAS J. SARGENT (1976) proposed a procedure which is directly derived from the Granger causality definition. It is usually denoted as the *direct Granger procedure*. LARRY D. HAUGH and DAVID A. PIERCE (1977) proposed a test which uses the estimated residuals of the univariate models for x and y. Finally, CHENG HSIAO (1979) proposed a procedure to identify and estimate bivariate models which – like the direct Granger procedure – is based on autoregressive representation and can also be interpreted (at least implicitly) as causality tests.

2.4.1 A Wald Test for Granger-Causality:

Let y_t be a VAR(p) process, partitioned into subprocesses z_t and x_t, that is, $y'_t = (z'_t, x'_t)$ we can define a causality from x_t to z_t and vice versa as in the previous section. Causality can be characterized by specific zero constraints on the VAR coefficients. Thus, in an estimated VAR(p) system, if we want to test for Granger-causality, we need to test zero constraints for the coefficients, we can derive the *asymptotic* tests of such constraints as:

More generally we consider testing, $C\beta = c \iff A_{jk,i} = 0$

$$H_0 : C\beta = c \qquad against \qquad H_1 : C\beta \neq c,$$

where C is an $(N \times (K^2p + K))$ matrix of $rankN$ and c is an $(N \times 1)$ vector. Assuming that

$$\sqrt{T}(C\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, \Gamma^{-1} \otimes \Sigma_u)$$

is an LS/ML estimation,we get

$$\sqrt{T}(C\hat{\beta} - C\beta) \xrightarrow{d} \mathcal{N}[0, C(\Gamma^{-1} \otimes \Sigma_u)C']$$

Hence the *Wald statistic* is given as:

$$T(C\hat{\beta} - c)'[C(\Gamma^{-1} \otimes \Sigma_u)C']^{-1}(C\hat{\beta} - c) \xrightarrow{d} \chi^2(N)$$

Replace Γ and Σ_u by their usual estimators $\hat{\Gamma} = ZZ'/T$ and $\widehat{\Sigma_u} = \frac{T}{T-K_p-1}\widetilde{\Sigma_u}$, the resulting statistic is:

$$\lambda_W = (C\hat{\beta} - c)'[C((ZZ')^{-1} \otimes \hat{\Sigma_u})C']^{-1}(C\hat{\beta} - c)$$

still has an asymptotic χ^2-distribution with N degrees of freedom, and also the condition $\frac{[C((ZZ')^{-1} \otimes \widetilde{\Sigma_u} C')]^{-1}}{T}$ is a consistent estimator of $[C(\Gamma^{-1} \otimes \Sigma_u)C']^{-1}$
Hence, we have the following result.
Proposition: *Asymptotic Distribution of the Wald Statistic* Suppose $\sqrt{T}(C\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, \Gamma^{-1} \otimes \Sigma_u)$ holds. Furthermore, $plim(ZZ'/T) = \Gamma, plim\widetilde{\Sigma_u} = \Sigma_u$ are both nonsingular and $H_0 : C\beta = c$ is true, with C being an $(N \times (K^2p + K))$ matrix of rank N. Then

$$\lambda_W = (C\hat{\beta} - c)'[C((ZZ')^{-1} \otimes \hat{\Sigma_u})C']^{-1}(C\hat{\beta} - c) \xrightarrow{d} \chi^2(N).$$

In practice, it may be useful to make adjustments to the statistic or the critical values of the test to compensate for the fact that the matrix $\Gamma^{-1} \otimes \Sigma_u$ is unknown and has been replaced by an estimator. Working in that direction, we note that

$$NF(N, T) \xrightarrow[T \to \infty]{d} \chi^2(N),$$

where $F(N, T)$ denotes an F random variable with N and T degrees of freedom (d.f.). Because an $F(N, T)$-distribution has a fatter tail than the $\chi^2(N)$-distribution divided by N, it seems reasonable to consider the test statistic $\lambda_F = \lambda_W/N$, in conjunction with critical values from some $F-$distribution. The question is then what numbers of degrees of freedom should be used? From the foregoing discussion it is plausible to use N as the numerator degrees of freedom. On the other hand, any sequence that goes to infinity with the sample size qualifies as a candidate for the denominator d.f. The usual F-statistic for a regression model with nonstochastic regressors has denominator d.f. equal to the sample size minus the number of estimated parameters. Assume that the vector $\mathbf{y}$ with KT observations and β contains $K(Kp+1)$ parameters. Hence, we have the approximate distributions $\lambda_F \approx F(N, KT - K^2p - K) \approx F(N, T - Kp - 1)$.

2.4.1.1 Causal analysis for bivariate VAR

For a bivariate system, y_t, x_t defined by

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} A_{11}(B) & A_{12}(B) \\ A_{21}(B) & A_{22}(B) \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_{xt} \end{bmatrix}$$

$$= \begin{bmatrix} \Phi_{11}(B) & \Phi_{12}(B) \\ \Phi_{21}(B) & \Phi_{22}(B) \end{bmatrix} \begin{bmatrix} u_{yt-1} \\ u_{xt-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_{xt} \end{bmatrix}$$

x_t does not Granger-cause y_t if $\Phi_{12}(B) = 0$ or $\Phi_{12,i} = 0$ for $i = 1, 2, 3, ...$ This condition is equivalent to $A_{12,i} = 0$ for $i = 1, 2, 3, ..., p$. In other words, this corresponds to the restrictions that all cross - lags coefficients are all zeros which can be tested by Wald statistics.

We now turn to determine the causal direction for bivariate VAR system. For ease of illustration, I shall focus on bivariate AR(1) process so that $A_{ij}(B) = A_{ij}; i, j = 1, 2$ as defined above. The results can be easily generalized to AR(p) case. Four possible causal directions between x and y are:

1. Feedback, $H_0, x \leftrightarrow y$ $$H_0 = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

2. Independent, $H_1 : x \perp y$ $$H_1 = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$$

3. x causes y but y does not cause x, $H_2 : y \not\rightarrow x$ $$H_2 = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

4. y causes x but x does not cause y, $H_3 : x \not\rightarrow y$ $$H_3 = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$$

Caines, Keng and Sethi(1981) proposed a two-stage testing procedure for determining causal directions. In the first stage, test $H_1(null)$ against H_0, $H_2(null)$ against H_0, and $H_3(null)$ against H_0. If necessary, test $H_1(null)$ against H_2, and $H_1(null)$ against $H3$. See Liang, Chou and Lin(1995) for an application.

2.4.1.2 Causal analysis for Vector ARMA model

Let y be $n \times 1$ stationary vector generated

$$\Phi(B)y_t = \Theta(B)a_t$$

y_i does not cause y_j if and only if

$$det(\Phi_i(z), \Theta_{(j)}(z)) = 0$$

where $\Phi_i(B)$ is the ith column of the matrix $\Phi(z)$ and $\Theta_{(j)}(z)$ is the matrix $\Theta(z)$ without its jth column. For bivariate (two-group) case,

$$\begin{bmatrix} \Phi_{11}(B) & \Phi_{12}(B) \\ \Phi_{21}(B) & \Phi_{22}(B) \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} \Theta_{11}(B) & \Theta_{12}(B) \\ \Theta_{21}(B) & \Theta_{22}(B) \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

Then, y_i does not cause y_j if and only if

$$\Phi_{21}(z) - \Theta_{21}(z)\Theta^{-1}(z)_{11}\Phi_{11}(z) = 0$$

If $n_1 = n_2 = 1$, then y_i does not cause y_j if and only if

$$\Theta_{11}(z)\Phi_{12}(z) - \Theta_{21}(z)\Phi_{11}(z) = 0$$

General testing procedures are:

1. Build a multivariate ARMA model for y_t,
2. Derive the noncausality conditions in term of AR and MA parameters, say $R_j(\beta_l) = 0, j = 1, ..., K$
3. Choose a test criterion, Wald, LM or LR test.

Let $T(\hat{\beta}_l) = (\frac{\partial R_j(B)}{\partial \beta_l}|_{\beta_l - \hat{\beta}_l})_{k\times k}$ Let $V(\beta_l)$ be the asymptotic covariance matrix of$\sqrt{N}(\beta_l \hat{=} \beta_l)$. Then the Wald and LR test statistics are:

$$\xi_w = NR(\hat{\beta}_l)'[T(\hat{\beta}_l)'V(\hat{\beta}_l)T(\hat{\beta}_l)]^{-1}R(\hat{\beta}_l),$$

$$\xi_{LR} = 2(L(\hat{\beta}, X) - L(\hat{\beta *}, y))$$

where $\hat{\beta *}$ isthe MLE of β under the constraint of noncausality.

To illustrate, let y_t be a invertible 2- dimensional ARMA(1,1) model.

$$\begin{bmatrix} 1-\phi_{11}B & -\phi_{12}B \\ -\phi_{21}B & 1-\phi_{22}B \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} 1-\theta_{11}B & \theta_{12}B \\ \theta_{21}B & \theta_{22}B \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

X_1 does not cause X_2 if and only if

$$\begin{aligned} \Theta_{11}(z)\Phi_{21}(z) - \Theta_{21}(z)\Phi_{11}(z) &= 0 \\ (\phi_{21} - \theta_{21})z + (\theta_{11}\theta_{21} - \phi_{21}\theta_{11})z^2 &= 0 \\ \phi_{21} - \theta_{21} &= 0 \qquad \phi_{11}\theta_{21} - \phi_{21}\theta_{11}) = 0 \end{aligned}$$

For the vector, $\beta_l = (\phi_{11}, \phi_{21}, \theta_{11}, \theta_{21})'$, the matrix

$$T(\beta_l) = \begin{pmatrix} 0 & \theta_{21} \\ 1 & -\theta_{11} \\ 0 & -\phi_{21} \\ -1 & \phi_{11} \end{pmatrix}$$

might not be nonsingular under the null of $H_0 : X_1$ does not cause X_2.

Remarks:

- The conditions are weaker than $\phi_{21} = \theta_{21} = 0$
- $\phi_{21} - \theta_{21} = 0$, is a necessary condition for H_0, $\phi_{21} = \theta_{21} = 0$ is sufficient condition and $\phi_{21} - \theta_{21} = 0$ and $\phi_{11} = \theta_{11}$are sufficient for H_0.

Let$H_0 : X_1$ does not cause X_2. Consider the following hypotheses:

$$\begin{aligned} H_0^1 &: \phi_{21} - \theta_{21} = 0; \\ H_0^2 &: \phi_{21} = \theta_{21} = 0; \\ H_0^3 &: \phi_{21} \neq 0, \phi_{21} - \theta_{21} = 0 \qquad and \qquad \phi_{11} - \theta_{11} = 0 \\ \tilde{H}_0^3 &: \phi_{11} - \theta_{11} = 0 \end{aligned}$$

Then, $H_0^3 = \tilde{H}_0^3 \cap H_0^1$, $H_0^2 \subseteq H_0 \subseteq H_0^1$, $H_0^3 \subseteq H_0 \subseteq H_0^1$.

Testing procedures:

1. Test H_0^1 at level α_1.If H_0^1 is rejected, then H_0 is rejected. Stop.
2. If H_0^1 is not rejected, test H_0^2 at level α_2. If H_0^2 is not rejected, H_0 cannot be rejected. Stop
3. If H_0^2 is rejected, test $\tilde{H}_0^3 : \phi_{11} - \theta_{11} = 0$ at level α_2. If $\tilde{H}_0^3$ is rejected, then H_0 is also rejected. If $\tilde{H}_0^3$ is not rejected, then H_0 is also not rejected.

2.4.2 Causal analysis for nonstationary processes

The asymptotic normal or χ^2 distribution in previous section is build upon the assumption that the underlying processes y_t is stationary. The existence of unit root and cointegration might make the traditional asymptotic inference invalid. Here, I shall briefly review unit root and cointegration and their relevance with testing causality. In essence, cointegration, causality test, VAR model and IR are closely related and should be considered jointly.

2.4.3 Integrated series

Determination of the integration order of our data is a step to exploring the statistical properties of the data. Integrated series are non stationary processes with the basic property of becoming stationary after differencing. Engle & Granger (1987) [17] defined a series to be integrated of order d (denoted $y_t \sim I(d)$) if it is stationary after differencing d times. Empirical studies have shown that macroeconomic series appear to be $I(1)$ as suggested by the typical spectral shape (Granger, 1996) [25], as by analysis of Box Jenkins (1970) modelling techniques or by direct testing as in Nelson & Plosser (1982). In the case of the variables we have considered for our analysis, in order to determine the order of integration of the series, the *Dickey Fulley test or Phillip Perron* test is performed.

2.4.3.1 Unit root:

What is *unit root?*
The time series y_t as defined in $A(B)y_t = C(B)\epsilon_t$ has an unit root if $A_p(1) = 0, C(1) \neq 0$
Why do we care about unit root?

- For y_t, the existence of unit roots implies that a shock in ϵ_t has permanent impacts on y_t.
- If y_t has a unit root, then the traditional asymptotic normality results usually no longer apply. We need different asymptotic theorems.

It is common to see macroeconomic and financial variables increase or decrease over time. This is due to the fact that, improvements of innovation and technology can lead to the increase of a country's over time. Since such variables show stochastic trend, they are assumed tpo be integrated of some order. Consider the simple AR(1) model:

$$y_t = \theta y_{t-1} + \epsilon_t \tag{2.3}$$

Determining the order of integration is equivalent to testing the null hypothesis that in the autoregressive model, $\theta = 1$. If the test holds, that is, the series are integrated. A pictorial estimation from the ACF and PACF plot can also give us similar information. Most often, difference stationary and trend stationary models of financial and economic time series imply different predictions and analysis. Deciding the model to use is therefore important for applied forecasters. Forecasters usually consider three choices to make on their data: always difference data, never difference or use a unit root pretest.

2.4.4 The Augumented Dickey- Fulley (ADF) test

The ADF test build on the original Dickey Fulley test for unit root. The Dickey Fulley test estimated by fitting the model with Ordinary Least Squares has the equation

$$labelar2\Delta y_t = (\delta - 1)y_{t-1} + \epsilon_t \cong \theta y_{t-1} + \epsilon_t \tag{2.4}$$

When there is an evidence of presence of serial correlation in the stationary residual, ϵ_t makes bias the result. *Said & Dickey (1984)* [49], modified equation 2.3 by adding lags to the autocorrelation. The parametric transformation of the model made captures the serial correlation in the residual. Depending on constant or trend component in the model, the ADF test consider these three different regression models.

$$\Delta y_t = \theta y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} \epsilon_t \tag{2.5}$$

$$\Delta y_t = \beta_1 + \theta y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} \epsilon_t \tag{2.6}$$

$$\Delta y_t = \beta_1 + \beta_2 t + \theta y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} \epsilon_t \tag{2.7}$$

The test statistic for the ADF test is given by $t_{\theta=0} = \frac{\hat{\theta}-1}{\sigma(\hat{\theta})}$ where $\hat{\theta}$ is the OLS estimate of θ and $\sigma\hat{\theta}$ is the standard error of the estimate. We note that $\theta = 0$ indicate existence of a unit root in equation 2.5. The t - statistic for θ is performed on the null hypothesis $H_0 : \theta = 0$ of unit root against the alternative hypothesis $(H_1 \neq 0)$ of stationarity. The test statistic here do not follow the student *t-distribution* since the test is performed on the residual rather the raw series.

2.4.5 The Phillip Perron (PP) test

In statistics, the Phillips - Perron test (named after Peter C. B. Phillips and Pierre Perron) is a unit root test [1]. That is, it is used in time series analysis to test the null hypothesis that a time series is integrated of order 1. It builds on the Dickey–Fuller test of the null hypothesis $\rho = 0$ in $\Delta y_t = \rho y_{t-1} + u_t$, where Δ is the first difference operator. Like the augmented Dickey - Fuller test, the Phillips - Perron test addresses the issue that the process generating data for y_t might have a higher order of autocorrelation than is admitted in the test equation - making y_{t-1} endogenous and thus invalidating

[1] see Phillips, P. C. B.; Perron, P. (1988) [39]

the Dickey - Fuller t-test. Whilst the augmented Dickey - Fuller test addresses this issue by introducing lags of Δy_t as regressors in the test equation, the Phillips–Perron test makes a non-parametric correction to the t-test statistic. The test is robust with respect to unspecified autocorrelation and heteroscedasticity in the disturbance process of the test equation.

Davidson and MacKinnon (2004) report that the Phillips–Perron test performs worse in finite samples than the augmented Dickey–Fuller test [2].

2.4.6 Cointegration:

What is *cointegration?*
When linear combination of two $I(1)$ process become an $I(0)$ process, then these two series are cointegrated.
Why do we care about cointegration?

- Cointegration implies existence of long-run equilibrium;
- Cointegration implies common stochastic trend;
- With cointegration, we can separate short- and long- run relationship among variables;
- Cointegration can be used to improve long-run forecast accuracy;
- Cointegration implies restrictions on the parameters and proper accounting of these restrictions could improve estimation efficiency.
- Cointegration introduces one additional causal channel (error correction term) for one variable to affect the other variables. Ignoring this additional channel will lead to invalid causal analysis.

Definition 11: Let x_t and y_t be two univariate discrete time processes integrated of order 1, that is $x_t \sim I(1)$, and $y_t \sim I(1)$. we say that x_t and y_t are cointegrated if

$$\lambda_1 x_t + \lambda_2 y_t \sim I(0)$$

where $\lambda_1, \lambda_2 \neq 0$

If $y_t = (y_{1t}, y_{2t}, ..., y_{kt})$ is an $I(1)$ process, we sat that the variables $y_{1t}, y_{2t}, ..., y_{kt}$ are cointegrated if

$$\lambda_1 y_{1t} + \lambda_2 y_{2t} + \cdots + \lambda_k y_{kt} \sim I(0)$$

where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_k)' \neq 0$

[2] see Davidson, Russell; MacKinnon, James G. (2004) [13]

We can test for cointegration between x_t and y_t by employing the ADF/PP test using the Engle - Granger two step procedure for the hypothesis.
$H_0 : \beta = 0$ (No cointegration exist)
$H_1 : \beta < 0$ (Cointegration exist)
The test rely on rejecting the null hypothesis in favor of cointegration between x_t and y_t and this require significant computed $t - values$ for β using the critical values reported in Mackinnon(1991).

Recommended procedures for testing cointegration:

1. Determine order of VAR(p). Suggestion choose the minimal p such that the residuals behave like vector white noise;

2. Determine type of deterministic terms: no intercept, intercept with constraint, intercept without constraint, time trend with constraint, time trend without constraint. Typically, model with intercept without constraint is preferred;

3. Use trace or λ_{max} tests to determine number of unit root;

4. Perform diagnostic checking of residuals;

5. Test for exclusion of variables in cointegration vector;

6. Test for weak erogeneity to determine if partial system is appropriate;

7. Test for stability;

8. Test for economic hypothesis that are converted to homogeneous restrictions on the cointegration vectors and/or loading factors.

2.4.6.1 Johansen cointegration rank test

In statistics, the Johansen test [3], is a procedure for testing cointegration of several, say k, $I(1)$ time series. This test permits more than one cointegrating relationship so is more generally applicable than the *Engle - Granger test* which is based on the Dickey - Fuller (or the augmented) test for unit roots in the residuals from a single (estimated) cointegrating relationship [4]

There are two types of Johansen test, either with trace or with eigenvalue, and the inferences might be a little bit different. The null hypothesis for the trace test is that the number of cointegration vectors is $r = r* < k$, vs. the alternative that $r = k$. Testing proceeds sequentially for $r* = 1, 2$,etc. and the first non - rejection of the null is taken as an estimate of r. The null hypothesis for the "maximum eigenvalue" test

[3] see Johansen, Søren (1991) [29]
[4] see Davidson, James (2000) [14]

is as for the trace test but the alternative is $r = r * +1$ and, again, testing proceeds sequentially for $r* = 1, 2, etc.$, with the first non - rejection used as an estimator for r.

The test make use of two likelihood ratio (LR) statistics; the trace statistic (λ_{trace}) and the maximum eigenvalue statistic (λ_{max}) which are all based on the estimated eigenvalue $\hat{\lambda}$ of Λ [5] The sequential test for the trace test rely on the hypothesis:
$H_0 : r = r_0$ against $H_1 : r_0 \leq r \leq K$ with LR statistic given by

$$\lambda_{tarce}(r) = -T \sum_{i=r_0+1}^{K} Ln(1 - \hat{\lambda}_i) \tag{2.8}$$

A similar test for the maximum eigenvalue however, considers a different alternative hypothesis that the rank $r = r_0 + 1$. The LR statistic here is given as:

$$\hat{\lambda}_{max}(r_0 + 1) = -TLn(1 - \hat{\lambda}_{r+1}) \tag{2.9}$$

Both tests can be used for the cointegration test. If the test statistic is greater than the critical value at the chosen significance level, the null hypothesis that exactly r_0 vectors are cointegrated is rejected. This is sequentially done until the null hypothesis cannot be rejected, consequently the r value at the null hypothesis becomes the accepted cointegration rank.

2.5 Vector Error Correction Model (VECM)

Suppose a unit root test of the VAR(p) process in section 4.2 shows that the variables are integrated (say $I(d)$) and the Johansen test indicate cointegration relation(s) among the variables, the vector error correction model becomes the appropriate model to capture the long term equilibrium relationship of the variables. A VAR(p) in difference to remove the unit root will eventually eliminate this long term relation from the trending stochastic variables and will not be an appropriate model.
Let the time series y_t be $k-$ dimensional VAR(p) series defined as

$$A(B)y_t = \epsilon_t \tag{2.10}$$

where B is a matrix of $r-$ cointegration vectors, such that, $A(B) = I - A_1B - A_2B^2 - \cdots - A_rB^r$. We can rewrite 2.10 in the Vector Error Correction Model (VECM), that is

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \epsilon_t \tag{2.11}$$

[5]By construction, the matrix $\hat{\Lambda}$ is symmetric positive definite and the eigenvalues are real and non - negative.

where $A(-1) = -\Pi$. If y_t. If the series is a cointegrated process then

$$-\Pi_{(K\times K)} = (\alpha)_{K\times r}(\beta)_{r\times K}, 0 < r < K$$

and so we have

$$\underbrace{\Delta y_t}_{I(0)} = \underbrace{-\alpha \underbrace{\beta' y_{t-1}}_{I(1)}}_{I(0)} + \underbrace{\sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i}}_{I(0)} + \underbrace{\epsilon_t}_{I(0)} \tag{2.12}$$

If $y_t \sim T(1)$ but not cointegrated, thenthis specification is correct, because in this case it is possible to show that Π.

Following our interest in modelling the Export (EXP), Gross Domestic Product (GDP) and Exchange Rate (EXC), we rewrite equation 2.11 to meet our model preference as follows:

$$\begin{aligned}
\Delta LGDP_t &= \Pi_1 ECT_{t-1} + \Gamma_{11}\Delta LGDP_{t-1} + \Gamma_{12}\Delta LEXP_{t-1} + \Gamma_{13}\Delta LEXC_{t-1} + \epsilon_{1t} \\
\Delta LEXP_t &= \Pi_2 ECT_{t-1} + \Gamma_{21}\Delta LEXP_{t-1} + \Gamma_{22}\Delta GDP_{t-1} + \Gamma_{23}\Delta LEXC_{t-1} + \epsilon_{2t} \\
\Delta LEXC_t &= \Pi_3 ECT_{t-1} + \Gamma_{31}\Delta LEXC_{t-1} + \Gamma_{32}\Delta LEXP_{t-1} + \Gamma_{33}\Delta LGDP_{t-1} + \epsilon_{3t}
\end{aligned}$$

where ECT is the error correction, herein y_{t-1}

2.5.0.1 Unit root, Cointegration and Causality

For a VAR system, X_t with possible unit root and cointegration, the usual causality test for the level variables could be misleading. Let $X_t = (X_{1t}, X_{2t}, X_{3t})'$ with n_1, n_2, n_3 dimension respectively. The VAR level model is:

$$\begin{aligned}
X_t &= J(B)X_{t-i} + u_t \\
&= \sum_{i=1}^{k} J_i X_{t-i} + u_t
\end{aligned}$$

Now the null hypothesis that X_3 does not cause X_1 can be formulated as:

$$H_0 : J_{1,13} = J_{2,13} = \cdots = J_{k,13} = 0$$

I will denote by F_{LS} the Wald Statistics for testing H_0

1. If X_t has unit root and is not cointegrated, F_{LS} converges to a limiting distribution which is the sum of χ^2 and unit root distribution. The test is similar and critical values can be constructed. Yet, it is more efficient and easier to difference X_t and

test causality for the differenced VAR.

2. If there issufficient cointegration for X_3 then$F_{LS} \to \chi^2_{n_1,n_2k}$. , More specifically, let $A = (A_1, A_2, A_3)$ be the cointegration vector and asssume that the usual asymptotic distribution results hold if $rank(A_3) = n_3$, ie. all X_3 appear in the cointegration vector.

3. If there is not sufficient cointegration, ie. not all X_3 appears in the cointegration vector, then the limiting distribution contain unit root and nuisance parameters.

For the error correction model,

$$\Delta X_t = J^*(B)\Delta X_{t-1} + \Gamma A' X_{t-1} + u_t$$

where Γ, A are the loading matrix and cointegration vector respectively. Partition Γ, A conforming toX_1, X_2, X_3. Then, if $rank(A_3) = n_3$ or $rank(\Gamma_1) = n_1$, $F_{ML} \to \chi^2_{n_1,n_3k}$., In other words, testing with ECM the usual asymptotic distribution hold when there are sufficient cointegrations or sufficient loading vector.

2.5.0.2 Toda and Yamamoto:

Toda and Yamamoto (1995) proposed a test of causality without pretesting cointegration. For an $VAR(p)$ process and each series is at most $I(k)$, then estimate the augmented $VAR(p+k)$ process even the last k coefficient matrix is zero.

$$X_t = A_1 X_{t-1} + \cdots + A_{p+k} X_{t-p-k} + U_t$$

and perform the usual Wald test $A_{kj,i} = 0, 1, 2, \cdots, p$. The test statistics is asymptotical χ^2 with degree of freedom m being the number of constraints. The result holds no matter whether X_t is $I(0)$ or $I(1)$ and whether there exist cointegration.
As there is no free lunch under the sun, the Toda-Yamamoto test suffer the following weakness.

- insufficient as compared with ECM where cointegration is explicitly considered.
- Cannot distinguish between short run and long run causality.
- Cannot test for hypothesis on long run equilibrium, say PPP which is formulated on cointegration vector.

One more remark: Cointegration between two variables implies existence of long-run causality for at least one direction. Testing cointegration and causality should be considered jointly.

2.5.1 Specific Tests:

2.5.1.1 The Direct Granger Procedure:

This procedure proposed by **T.J. SARGENT (1976)** is derived directly from the Granger definition of causality. Similar to the method of **C.W.J. GRANGER (1969)**, a linear prediction function is employed. In the following, let x and y be two stationary variables. To test for simple causality from x to y, it is examined whether the lagged values of x in the regression of y on lagged values of x and y significantly reduce the error variance. By using OLS, the following equation is estimated:

$$y_t = \alpha_0 + \sum_{k=1}^{k_1} \alpha_{11}^k y_{t-k} + \sum_{k=k_0}^{k_2} \alpha_{12}^k x_{t-k} + u_{1,t} \qquad (*)$$

with $k_0 = 1$. An F test is applied to test the null hypothesis, $H_0 : \alpha_{12}^1 = \alpha_{12}^2 = \cdots = \alpha_{12}^{k_2}$ By changing x and y in (*), it can be tested whether a simple causal relation from y to x exists. There is a feedback relation if the null hypothesis is rejected in both directions. To test whether there is instantaneous causality we finally set $k_0 = 0$ in relation (*) and perform a t or F test for the null hypothesis $H_0 : \alpha_{12}^1 = 0$. Accordingly, the correspondding null hypothesis can be tested for x. According to the *Theorem* given above, we expect the same result for testing the equation for y and for$x$. However, as our data are based on finite samples, we will generally get different numerical values for the test statistics. However, with $k_1 = k_2$, i.e. if we include the same number of lagged variables for the dependent as well as for the explanatory variable in both test equations, we get exactly the same numerical values for the test statistics. The reason for this is that the t or F statistics are functions of the partial correlation coefficient between x and y. Its value does not depend on the direction of the regression; it only depends on the correlation between the two variables and the set of conditioning variables which are included. If $k_1 = k_2$, the same conditioning variables are included irrespectively of the dependent variable.

One problem with this test is that the results are strongly dependent on the number of lags of the explanatory variable, k_2. There is a trade-off: the more lagged values we include, the better the influence of this variable can be captured. This argues for a high maximal lag. On the other hand, the power of this test is the lower the more lagged values are included. Two procedures have been developed to solve this problem. In general, different values of k_2 (and possibly also of k_1) are used to inspect the sensitivity of the results to the number of lagged variables. If we include an explanatory variable, the number of estimated parameters, m, has to be adjusted. If, besides the constant term on the right hand side, we include k_1 lagged values of the dependent and k_2 values of additional variables, it holds that $m = k_1 + k_2 + 1$.

2.5.1.2 The Haugh-Pierce Test:

This procedure which was first proposed by **L.D. HAUGH (1976)** and later on by **L.D. HAUGH and D.A. PIERCE (1977)** is based on the crosscorrelations $\rho_{ab}(k)$ between the residuals a and b of the univariate ARMA models for x and y. In a first step, these models have to be estimated. By using the Box-Pierce Q statistic (or the Box-Ljung Q statistic) it is checked whether the null hypothesis – that the estimated residuals are white noise – cannot be rejected. Then, analogous to the Q statistic, the following statistic is calculated:

$$S = T \cdot \sum_{k=k_1}^{k_2} \hat{\rho}^2_{ab}(k) \qquad (**)$$

Under the null hypothesis $H0 : \rho_{ab}(k) = 0$ for all k with $k_1 \leq k \leq k_2$, this statistic is asymptotically χ^2 disdributed with $k_2 - k_1 + 1$ degrees of freedom. It can be checked for $k_1 < 0 \wedge k_2 > 0$ whether there is any causal relationship at all. If this hypothesis can be rejected, it can be checked for $k_1 = 1 \wedge k_2 \geq 1$ whether there is a simple causal relation from x to y. In the reverse direction, for $k_1 \leq -1 \wedge k_2 = -1$, it can be checked whether there is a simple causal relation from y to x. Finally, it can be tested by using $\rho_{ab}(0)$ whether there exists an instantaneous relation. However, the results of the last test are questionable as long as the existence of a feedback relation cannot be excluded.

But this is not the only problem that might arise with this procedure. **G. WILLIAM SCHWERT (1979)** showed that the power of this procedure, which uses correlations, is smaller than the power of the direct Granger procedure which uses regressions. Thus, following a remark by **EDGAR L. FEIGE and DOUGLAS K. PEARCE (1979)**, this test might only be a first step to analyse causal relations between time series. On the other hand, information on the relations between two time series, which is contained in crosscorrelations, can be useful even if no formal test is applied. This information offers a deeper insight into causal relations than just looking at the F and t statistics of the direct Granger procedure.

2.5.1.3 The Hsiao Procedure:

The procedure for identifying and estimating bivariate time series models proposed by **CHENG HSIAO (1979)** initially corresponds to the application of the direct Granger procedure. However, the lag lengths are determined with an information criterion. C. HSIAO proposed the use of the final prediction error. Again, the precondition is that the two variables are weakly stationary. The procedure is divided into six steps:

1. First, the optimal lag length k_1^* of the univariate autoregressive process of y is determined.

2. In a second step, by fixing k_1^* , the optimal lag length k_2^* of the explanatory variable x in the equation of y is determined.

3. Then k_2^* is fixed and the optimal lag length of the dependent variable y is again determined: $\bar{k}_1^*$.

4. If the value of the information criterion applied in the third step is smaller than that of the first step, x has a significant impact on y. Otherwise, the univariate representation of y is used. Thus, we get a (preliminary) model of y.

5. Steps (1) to (4) are repeated by exchanging the variables x and y Thus, we get a (preliminary) model for x.

6. The last step is to estimate the two models specified in steps (1) to (5) simultaneously to take into account the possible correlation between their residuals. Usually, the procedure to estimate *seemingly unrelated regressions* (SUR) developed by ARNOLD ZELLNER (1962) is applied.

The Hsiao procedure only captures the simple causal relations between the two variables. The possible instantaneous relation is reflected by the correlation between the residuals. However, by making theoretical assumptions about the direction of the instantaneous relation, it is possible to take into account the instantaneous relation in the model for y or in the model for x.

Chapter 3

Panel Data

3.1 Introduction

Before now, we have been concentrating on time series data and the use of appropriate procedures to estimate models using such data, especially when the data may be stationary or non-stationary (i.e., contain a unit root(s)). However, panel data (i.e., cross-sectional time series data with $i = 1, ..., N$ individuals' in each time period and with $t = 1, ..., T$ observations for each individual over time) are increasingly being used in both macro- as well as the more traditional microlevel studies of economic problems. At the macro-level there is increasing use of cross - country data to study such topics as purchasing power parity (Pedroni, 2001 and growth convergence (McCoskey, 2002 as well as familiar issues such as whether real gross domestic product data contain unit roots (Rapach, 2002 Micro-based panel data (such as those generated by national household surveys or surveys of firms) are also widely used where typically the data comprise large N and small T. Baltagi (2001) considers some of the major advantages (as well as limitations) of using panel data, such as how they allow for heterogeneity in individuals, firms, regions and countries, which is absent when using aggregated time series data. They also give more variability, which often leads to less collinearity among variables, while cross sections of time series provide more degrees of freedom and more efficiency (more reliable parameter results) when estimating models. The dynamics of adjustment are better handled using panels especially in micro-based studies involving individuals, and more complicated models can be considered involving fewer restrictions. The limitations of panel data are usually related to the design and collection of such information: not just missing data (e.g., from non - response) but also measurement errors, attrition in the panel over time and selectivity problems (including issues such as the weighting of data that is sampled on the basis of a particular stratification of the population). Model estimation using unbalanced panels (where there are not T observations on all i individuals in the data set) is more complicated, but often necessary given the impacts of the problems just outlined.

This chapter begins with a brief overview of econometric techniques for use with panel data (see Baltagi, 2001 for an extensive review of the area; and Baltagi, Fomby and Hill, 2000 for issues related specifically to the use of nonstationary panel data). It then considers the various approaches that have become popular for testing for Granger causality in panel data.

3.2 Panel Data:

3.2.1 Definition:

Panel data, also called longitudinal data or cross-sectional time series data, are data where multiple cases (people, firms, countries etc) were observed at two or more time periods. An example is the National Longitudinal Survey of Youth, where a nationally representative sample of young people were each surveyed repeatedly over multiple years. A panel data set contains repeated observations over the same units (individuals, households, firms), collected over a number of periods.
NB:

There are two kinds of information in cross-sectional time-series data: the cross-sectional information reflected in the differences between subjects, and the time-series or within-subject information reflected in the changes within subjects over time. Panel data regression techniques allow you to take advantage of these different types of information.

Eventhough panel data are typically collected at the micro - economic level, it gas now become more practical to gather individual time series of a number of countries or industries and analyse them simultaneously. This availability of repeated observations on the same units permits economists to specify and estimate more complicated and more realistic models than a single cross - section or a single time series would do. The disadvantages are more of a practical nature: because we repeatedly observe the same units, it is usually no longer appropriate to assume that different observations are independent. This may complicate the analysis, particularly in nonlinear and dynamic models. Also panel data sets may suffer from missing obswrvations in which case we need to adjust our analysis.

3.2.2 Panel Data and Modelling Techniques

In general , we could specify a linear model as

$$y_{it} = x'_{it}\beta_{it} + \epsilon_{it}, \qquad \forall i = 1, ..., N, \quad \forall t = 1, ..., T \tag{3.1}$$

where β_{it} measures the partial effects of x_{it} in period t for unit i. Of course, this model is much too general to be useful, and we need to put more structure on the coefficients β_{it}. The standard assumption, used in many empirical cases, is that β_{it} is constant for all i and t, except, possibly, the intercept term. This could be written as

$$y_{it} = \alpha_i + x'_{it}\beta + \epsilon_{it}, \qquad \forall i = 1, ..., N, \quad \forall t = 1, ..., T \tag{3.2}$$

where x_{it} is a K- dimensional vector of explanatory variables, not including a constant (also β are indexed from β_1 to β_k). This means that the effects of a change in x are the same for all units and all periods, but that the average level for unit i may be different from that for unit j. The α_i thus capture the effects of those variables that are peculiar to the $i-th$ individual and that are constant over time. In the standard case, ϵ_{it} is assumed to be independent and identically distributed over individuals and time, with mean zero and variance σ^2_ϵ. If we treat the α_i as N fixed unknown parameters, the model in 3.2 is referred to as the **standard fixed effects model**.

An alternative approach assumes that the intercepts of the individuals are different but that they can be treated as drawings from a distribution with mean μ and variance σ^2_α. The essential assumption here is that these drawings are independent of the explanatory variables in x_{it} (see below). This leads to the **random effects model**, where the individual effects α_i are treated as random. The error term in this model consists of two components: a time - invariant component α_i (In the random effects model, the α_i's are redefined to have a zero mean) and a remainder component ϵ_{it} that is uncorrelated over time. It can be written as

$$y_{it} = \mu + x'_{it}\beta + \alpha_i + \epsilon_{it}, \qquad \forall i = 1, ..., N, \quad \forall t = 1, ..., T \tag{3.3}$$

where μ denotes the intercept term, y_{it} is the dependent variable, x'_{it} is a K - dimensional row vector of time - varying explanatory variables and, β is a K - dimensional column vector of parameters, α_i is an individua l- specific effect and ϵ_{it} is an idiosyncratic error term.

NB: The possibility of treating the α_i's as fixed parameters has some great advantages, but also some disadvantages. Most panel data models are estimated under either the fixed effects or the random effects assumption.
We will assume throughout this chapter that each individual i is observed in all time periods t. This is a so - called balanced panel. The T observations for individual i can

be summarized as:

$$y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{it} \\ \vdots \\ y_{iT} \end{bmatrix}_{T\times 1}, X_i = \begin{bmatrix} x'_{i1} \\ \vdots \\ x'_{it} \\ \vdots \\ x'_{iT} \end{bmatrix}_{T\times K} \quad \epsilon_i = \begin{bmatrix} \epsilon'_{i1} \\ \vdots \\ \epsilon'_{it} \\ \vdots \\ \epsilon'_{iT} \end{bmatrix}_{T\times 1}$$

and NT observations for all individuals and time periods as

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{bmatrix}_{NT\times 1}, X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{NT\times K} \quad \epsilon = \begin{bmatrix} \epsilon'_1 \\ \vdots \\ \epsilon'_i \\ \vdots \\ \epsilon'_N \end{bmatrix}_{NT\times 1}$$

The data generation process (dgp) is described by:

Linearity

$y_{it} = \mu + x'_{it}\beta + \alpha_i + \epsilon_{it}$ where $E[u_{it}] = 0$ and $E[\alpha_i] = 0$. The model is linear in terms of parameters μ, β, effect α_i, and error ϵ_{it}.

Independence:

$\{X_i, y_i\}_{i=1}^N$ i.i.d. (independent and identically distributed). The observations are independent across individuals but not necessarily across time. This is guaranteed by random sampling of individuals.

Strict Exogeneity:

$E[\epsilon_{it}|X_i, \alpha_i] = 0$ (mean independent). The idiosyncratic error term ϵ_{it} is assumed uncorrelated with the explanatory variables of all past, current and future time periods of the same individual. This is a strong assumption which e.g. rules out lagged dependent variables. It is also assumed that the idiosyncratic error is uncorrelated with the individual specific effect.

Error Variance:

- $V[\epsilon_{it}|X_i, \alpha_i] = \sigma^2_\epsilon I, \quad \sigma^2_\epsilon > 0$, and finite, (homoscedastic and no serial correlation)

- $V[\epsilon_{it}|X_i, \alpha_i] = \sigma^2_{\epsilon,it} > 0$, finite and $Cov[\epsilon_{it}, \epsilon_{is}|X_i, \alpha_i] = 0; \forall s \neq t$ (no serial correlation)

- $V[\epsilon_{it}|X_i, \alpha_i] = \Omega_{\epsilon,i}(X_i, \alpha_i)$; is positive definite and finite

3.2.2.1 The Random Effects Model

It is commonly assumed in regression analysis that all factors that affect the dependent variable, but that have not been included as regressors, can be appropriately summarized by a random error term. In this case, this leads to the assumption that the α are random factors, independently and identically distributed over individuals. Thus we write the random effects model as

$$y_{it} = \mu + x'_{it}\beta + \alpha_i + \epsilon_{it}, \qquad \epsilon_{it} \sim i.i.d(0, \sigma^2_\epsilon), \qquad \alpha_i \sim i.i.d(0, \sigma^2_\alpha) \tag{3.4}$$

where $\alpha + \epsilon_{it}$ is treated as an error term consisting of two components: an individual specific component, which does not vary over time, and a remainder component, which is assumed to be uncorrelated over time. That is, all correlation of the error terms over time is attributed to the individual effects α_i. It is assumed that α_i and ϵ_{it} are mutually independent and independent of x_{js} (for all j and s). This implies that the OLS estimator for μ and β from 3.4 is unbiased and consistent. The error components structure implies that the composite error term $\alpha_i + \epsilon_{it}$ exhibits a particular form of autocorrelation (unless $\sigma^2_\alpha = 0$).

In the random effects model, the individual - specific effect is a random variable that is uncorrelated with the explanatory variables with the following assumptions:

- Unrelated Effects
  $E[\alpha_i|X_i] = 0$. This assumes that the individual-specific effect is a random variable that is uncorrelated with the explanatory variables of all past, current and future time periods of the same individual.

- Effect Variance
  a) $V[\alpha_i|X_i] = \sigma^2_\alpha < \infty$ (homoscedastic), this assumes constant variance of the individual specific effect.
  b) $V[\alpha_i|X_i] = \sigma^2_{\alpha,i}(X_i) < \infty$ (heteroscedastic)

- Identifiability
  a)$rank(X) = K+1 < NT$ and $E[X'_iX_i] = Q_{XX}$ is positive definite and finite. The typical element $w'_{it} = [1x'_{it}]$. b) $rank(X) = K + 1 < NT$ and $E[X'_i\Omega^{-1}_{\epsilon,i}] = Q_{XOX}$ is positive definite and finite.$\Omega_{\epsilon,i}$ is defined below.These assume that the regressors including a constant are not perfectly collinear, that all regressors (but the constant) have non-zero variance and not too many extreme values.

where

$$\Omega_\epsilon = V[\epsilon_i|X] = \begin{bmatrix} \Omega_{\epsilon,1} & \dots & 0 & \dots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & & \Omega_{\epsilon,i} & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & 0 & \dots & \Omega_{\epsilon,N} \end{bmatrix}_{NT\times NT}$$

with typical element

$$\Omega_{\epsilon,i} = V[\epsilon_i|X_i] = \begin{bmatrix} \sigma_\epsilon^2 & \dots & \sigma_\alpha^2 & \dots & \sigma_\alpha^2 \\ \vdots & \ddots & & & \vdots \\ \sigma_\alpha^2 & & \sigma_\epsilon^2 & & \sigma_\alpha^2 \\ \vdots & & & \ddots & \vdots \\ \sigma_\alpha^2 & \dots & \sigma_\alpha^2 & \dots & \sigma_\epsilon^2 \end{bmatrix}_{T\times T}$$

where $\sigma_\mu^2 = \sigma_\alpha^2 + \sigma_\epsilon^2$.

3.2.2.2 The Fixed Effects Model

The fixed effects model is simply a linear regression model in which the intercept terms vary over the individual units i, i.e.

$$y_{it} = \alpha_i + x_{it}'\beta + \epsilon_{it}, \qquad \epsilon_{it} \sim i.i.d(0, \sigma_\epsilon^2) \tag{3.5}$$

where it is usually assumed that all x_{it} are independent of all ϵ_{it}. We can write this in the usual regression framework by including a dummy variable for each unit i in the model. That is,

$$y_{it} = \sum_{j=1}^{N} \alpha_j d_{ij} + x_{it}'\beta + \epsilon_{it}, \tag{3.6}$$

where $d_{ij} = 1$ if $i = j$ and 0 elsewhere. We thus have a set of N dummy variables in the model. The parameters $\alpha_1, ..., \alpha_N$ and β can be estimated by ordinary least squares in 3.6. The implied estimator for β is referred to as the **least squares dummy variable (LSDV) estimator**

In the fixed effects model, the individual- specific effect is a random variable that is allowed to be correlated with the explanatory variables with the following assumptions.

- Related Effects

  $E[\alpha_i|X_i] \neq 0$. This assumes that the individual - specific effect is a random variable that is correlated with the explanatory variables of all past, current and future time periods of the same individual.

- absence of Variance Effects

- Identifiability
$(\ddot{X}) = K < NT$ and $E(\ddot{x}_i'\ddot{x}_i)$ is positive definite and finite, where the typical element $\ddot{x}_{it} = x_{it} - \bar{x}_i$ and $\bar{x}_i = \frac{1}{T}\sum_t x_{it}$. We assume that the time - varying explanatory variables are not perfectly collinear, that they have non - zero within - variance (i.e. variation over time for a given individual) and not too many extreme values. Hence, x_{it} cannot include a constant or any time-invariant variables. Note that only the parameters β but neither α is identifiable in the fixed effects model

3.3 Granger Causality in Panel Data

Let us consider two covariance statiomnary variables, denoted by x and y, observed on T periods and on N individuals. The theoretical framework generally used to test causality in panel data, is directly derived from the vectorial autoregressive representation proposed by *Holtz-Eakin and al.(1988).* For each individual $i = 1, .., N$, at time $t = 1, .., T$, we consider the following linear model:

$$y_{i,t} = \sum_{k=1}^{K} \gamma_i^{(k)} y_{i,t-k} + \sum_{k=0}^{K} \beta_i^{(k)} x_{i,t-k} + v_{i,t}, \tag{3.7}$$

with $K \in \mathcal{N}$ and $v_{i,t} = \alpha_i + \epsilon_{i,t}$, where $\epsilon_{i,t}$ are i.i.d. $(0, \sigma_\epsilon^2)$. For simplicity, individual effects α_i are supposed to be fixed. Initial conditions $(y_{i,-K}, ..., y_{i,0})$ and $(x_{i,-K}, ..., x_{i,0})$ of both individual processes $y_{i,t}$ and $x_{i,t}$ are given and observable. We assume that lag orders K are identical for all cross-section units of the panel and the panel is balanced. In a first part, we allow for autoregressive parameters $\gamma_i^{(k)}$ and regression coefficients slopes $\beta_i^{(k)}$ to differ across groups. However, contrary to *Weinhold (1996)* and *Nair-Reichert* and *Weinhold (2001)*, we assume that the autoregressive coefficients $\gamma^{(k)}$ and the regression coefficients slopes $\beta_i^{(k)}$ are constant. So, the model 3.7 is not a random coefficient model as in *Swamy (1970)*: it is a fixed coefficients model with fixed individual effects.

3.3.1 Assumptions:

In order to implement a simple Granger causality procedure of test, we consider the following assumptions.

Assumption (A1)

For each cross section unit $i = 1, .., N$, individual residuals$\epsilon_{i,t}$, $\forall t = 1, ..., T$, are independently and normally distributed with $E(\epsilon_{i,t} = 0)$ and $E(\epsilon_{i,t}^2) = \sigma_{\epsilon,i}^2$.

Assumption (A2)

Individual residuals $\epsilon_i = (\epsilon_{i,1}, ..., \epsilon i, T)'$, are independently distributed across groups: $E(\epsilon_{i,t}\epsilon_{j,s}) = 0, \forall i \neq j$ and $\forall(t, s)$.

Assumption (A3)

Both individual variables $x_i = (x_{i,1}, ..., x_{i,T})'$ and$y_i = (y_{i,1}, ..., y_{i,T})'$, are covariance stationary with $E(y_{i,t}^2) < \infty, E(x_{i,t}^2) < \infty, E(y_{i,t}y_{j,z}), E(x_{i,t}x_{j,z})$ and $E(y_{i,t}x_{j,z})$ are only function of the difference t - z, whereas $E(x_{i,t})$ and $E(y_{i,t})$ are independent of t.

The normality assumption in **A1** could be relaxed without difficulty when we derive the asymptotic distributions of our statistics of causality tests. On the contrary, the assumption **A2** which prevent from the possibility of serially correlated errors, possibly with different serial correlations across groups, is more crucial as we will see later. This simple two variables model constitutes the basic framework to study the Granger causality in a panel data context. The introduction of a panel data dimension allows to use both cross-sectional and time series informations to test the causality relationships between y and x. In particular, it leads to give the researcher a large number of observations, increasing the degree of freedom and reducing the collinearity among explanatory variables. So, it noticeably improves the efficiency of econometric tests of the Granger causality hypothesis. However, the use of panel data raises the issue of the heterogeneity of the causality relationships.

3.3.2 Causality and Heterogeneity:

The standard causality tests consist in testing linear restrictions on parameters$\beta_i = (\beta_i^{(1)}, ..., \beta_i^{(K)})'$. However, if panel data are used to test causality, we must be very careful of the issue of heterogeneity between individuals. The first source of heterogeneity is standard and comes from permanent cross sectional disparities between individuals. A pooled regression ignoring heterogenous intercepts, leads to a bias of the slope estimates γ_i and β_i and then could lead to fallacious inference in causality tests. Such heterogeneity is controlled by the introduction of individual effects α_i in model 3.7. The second source, which is more crucial, is related to the heterogeneity of the parameters $\beta_i^{(K)}$. This kind of heterogeneity directly affects the paradigm of the representative agent and so, the conclusions about causality relationships. It is well known that the estimates of autoregressive parameters β_i get under the wrong hypothesis $\beta_i = \beta_j, \forall(i, j)$ are biased (see Pesaran Smith 1995 for an AR (1) process). Then, if we impose the homogeneity of coefficients $\beta_i^{(K)}$, the statistics of causality tests can lead to a fallacious inference. Intuitively, the estimate $\hat{\beta}$ obtained in an homogeneous model will converge to a value close to the average of the true coefficients β_i, and that if this mean is itself close to zero, we risk to accept at wrong the hypothesis of no causality.

Note:

Beyond these statistical stakes, it is evident that an homogeneous specification of the relation between the variables x and y does not allow to give some interpretation of the relations of causality as soon as at least one individual of the sample has an economic behavior different from that of the others. For example, let us assume that there exists a relation of causality for a set of N countries, for which parameters $\beta_i^{(K)}$ are strictly identical. If we introduce into the sample, a set of N_1 countries for which, on the contrary, there is no relation of causality, what are the conclusions? Statistically, we can show that standard tests in homogeneous model, would then lead to conclude to accept or to reject the hypothesis of global causality, according to the value of the ratio $\frac{N}{N_1}$. However, these conclusions do not correspond to economic reality, since the classic paradigm of the representative agent does not hold. If we ignore this heterogeneity, the test of the causality hypothesis is nonsensical and may lead to a wrong conclusion according the relative size of the two subgroups.

For these reasons, several definitions will be proposed for the causality relationships that could occur in models with fixed coefficients. These definitions are based on the heterogeneity of the underlying processes. In model 3.7, under assumptions **A1**, four principal cases is considered. Let us define $E(y_{i,t}|\bar{y}_{i,t}, \bar{\bar{x}}_{i,t})$ the best linear predictor of $y_{i,t}$ given the set of past values of $y_{i,t}$, denoted $\bar{y}_{i,t} = (y_{i,-p}, ..., y_{i,0}, ..., y_{i,t-1})'$, and the set of past and present values of $x_{i,t}$, denoted $\bar{\bar{x}}_{i,t} = (x_{i,-p}, ..., x_{i,0}, ..., x_{i,t-1}, x_{i,t})'$. For simplicity, we assume that individual effects are fixed.

1. The first case corresponds to the Homogenous Non Causality (HNC) hypothesis. Conditionally to the specific error components of the model, this hypothesis implies that there does not exist any individual causality relationships:

$$E(y_{i,t}|\bar{y}_{i,t}) = E(y_{i,t}|\bar{y}_{i,t}, \bar{x}_{i,t}), \quad \forall i = 1, ..., N \tag{3.8}$$

   This de finition can be extended to the homogenous instantaneous non causality hypothesis as follows:

$$E(y_{i,t}|\bar{y}_{i,t}) = E(y_{i,t}|\bar{y}_{i,t}, \bar{\bar{x}}_{i,t}), \quad \forall i = 1, ..., N \tag{3.9}$$

2. The second case corresponds to the Homogenous Causality (HC) hypothesis, in which there exists N causality relationships:

$$E(y_{i,t}|\bar{y}_{i,t}) \neq E(y_{i,t}|\bar{y}_{i,t}, \bar{x}_{i,t}), \quad \forall i = 1, ..., N \tag{3.10}$$

   In this case, we assume that the N individual predictors, obtained conditionally to the same past values $\bar{x}_t = \bar{x}_{i,t} = \bar{x}_{j,t}$ and $\bar{y}_t = \bar{y}_{i,t} = \bar{y}_{j,t}$, are identical:

$$E(y_{i,t}|\bar{y}_t, \bar{x}_t) = E(y_{j,t}|\bar{y}_t, \bar{x}_t), \quad \forall i = 1, ..., N \tag{3.11}$$

The instantaneous homogenous causality hypothesis is then de fined by:

$$E(y_{i,t}|\bar{y}_{i,t}) \neq E(y_{i,t}|\bar{y}_{i,t}, \bar{\bar{x}}_{i,t}), \quad \forall i = 1, ..., N \tag{3.12}$$

$$E(y_{i,t}|\bar{y}_t, \bar{\bar{x}}_t) = E(y_{j,t}|\bar{y}_t, \bar{\bar{x}}_t), \quad \forall (i, j) \tag{3.13}$$

3. The third case corresponds to the HEterogenous Causality (HEC) hypothesis. Under HEC hypothesis, we assume first that there exists a causality relationships for all individual and second that there exists at least two individual for which the conditional mean of y given the same past values $\bar{y}_t$ and $\bar{x}_t$ are not identical.

$$E(y_{i,t}|\bar{y}_{i,t}) \neq E(y_{i,t}|\bar{y}_{i,t}, \bar{x}_{i,t}), \quad \forall i = 1, ..., N \tag{3.14}$$

$$E(y_{i,t}|\bar{y}_t, \bar{x}_t) \neq E(y_{j,t}|\bar{y}_t, \bar{x}_t), \quad \exists (i, j) \tag{3.15}$$

The corresponding instantaneous hypothesis are then obtained by substituting in these de finitions $\bar{x}_t$ by $\bar{\bar{x}}_t$.

4. The last case corresponds to the HEterogenous Non Causality (HENC) hypothesis. In this case, we assume that there exists at least one and at the most $N - 1$ equalities of the form:

$$E(y_{i,t}|\bar{y}_{i,t}) = E(y_{i,t}|\bar{y}_{i,t}, \bar{x}_{i,t}), \quad \forall i = 1, ..., N_1 \tag{3.16}$$

In other words, there exists at least one individual and at the most $N - 1$ for which there is no causality from x to y. The size of this subgroup is denoted N_1, with $N_1 < N$. The corresponding instantaneous hypothesis are then obtained by substituting in these de finitions $\bar{x}_{i,t}$ by $\bar{\bar{x}}_{i,t}$.

To sum up, we propose here to distinguish between the heterogeneity of the data generating process (DGP) and the heterogeneity of the causality relationship from x to y. In the HNC hypothesis, there does not exist any individual causality from x to y. On the contrary, in the HC and HEC cases, there is a causality relationships for each individual of the sample. In the HC case, the DGP is homogenous, whereas it is not the case in the HEC hypothesis. Finally in the $HENC$ hypothesis, there is an heterogeneity of the causality relationships since there is a subgroup of N_1 units for which the variable x does not cause y. We now propose a nested tests procedure to characterize these various causality relationships given the heterogeneity of the data generating process.

3.4 Testing Procedures:

If we consider the model 3.7, the general definitions of Granger causality from x to y imply to test linear restrictions on the parameters $\beta_i^{(k)}$ associated to the lagged variables $x_{i,t-k}$. The test procedure has three main steps.

The first step of the procedure consists in testing the homogeneity of the parameters (except the individual effects) of the VAR representation. Here, we only consider the homogeneity of the interest parameters $\beta_i^{(k)}$ and not the autoregressive parameters $\gamma_i^{(k)}$: under the null and the alternative hypothesis, we allow $\gamma_i^{(k)}$ to vary across cross sections. If the homogeneity hypothesis is accepted, we can test the Granger non causality hypothesis as in *Holtz-Eakin, Newey and Rosen (1988).* The null hypothesis of the Homogenous Non Causality (HNC) test is then defined by the nullity of all the common parameters $\beta^{(k)}$, for all the considered lags $k = 1, .., K$. If the null is accepted, the variable x does not Granger cause the variable y for all the individuals of the panel (HNC hypothesis). If the null is rejected, the variable x Granger causes the variable y, and the improvement of the forecasts on y is similar for the individual of the panel (HC hypothesis). Under the homogeneity hypothesis, the Homogenous Non Causality (HNC) test is then very similar to the unit root test proposed by *Levin and Lin (1992)*, since under the alternative hypothesis, the parameters are restricted to be homogenous across all units of the panel.
On the contrary, if we reject the homogeneity hypothesis, we propose to test the Homogenous Non Causality test against an alternative in which there is a subgroup of N_1 units with no causality relations and a subgroup of $N - N_1$ units for which x Granger causes y. The structure of this test is very similar to the unit root test in heterogeneous panels proposed by *Im, Pesaran and Shin (2002).* If the null is accepted, we conclude to the Homogenous Non Causality (HNC) hypothesis. If H_0 is rejected, two cases can be identified : if the dimension of the first subgroup is null ($N_1 = 0$), we get the HEterogenous Causality (HEC) case. If this dimension is strictly positive, the variable x Granger causes y for all units, but the dynamic relations between both variables is heterogenous: this is the HEterogenous Non Causality (HENC) case. The last test is then a test on the dimension N_1

3.4.1 Homogeneity Test

We consider a sample of N cross sections observed over T time periods, where T denotes the dimension for estimation after adjustment for the initial values. The first step of our procedure consists in testing the homogeneity of the regression slope coefficients associated to $x_{i,t-k}$ for all lags $k = 1, .., K$. In model 3.7, the corresponding test is de fined by:

$$H_0 : \beta_i^{(k)} = \beta_j^{(k)}, \quad \forall (i,j), \quad \forall k = 1, ..., K \tag{3.17}$$

This test is similar to standard homogeneity tests *(Hsiao 1986)*. In order to test these $(N-1)K$ linear restrictions, we compute the following F_h statistic:

$$F_H = \frac{\frac{(RSS_0 - RSS_1)}{K(N-1)}}{\frac{RSS_1}{N(T-2K-1)}} \tag{3.18}$$

where RSS_0 denotes the restricted sum of squared residual obtained under H_0 and RSS_1 corresponds to the residual sum of squares of the model 3.7 without any restriction. Under assumptions H_1 and H_2, if we assume that individual effects α_i are fixed, RSS_0 is given by the residual sum of squares obtained from the MLE which corresponds in this case to the *Within estimator*. If we stack the T periods observations for the i^{th} individual's characteristics into T elements columns, as:

$$y_i^{(-k)} = \begin{bmatrix} y_{i,-k+1} \\ \vdots \\ y_{i,i} \\ \vdots \\ y_{i,T-k} \end{bmatrix}_{(T\times 1)} \qquad x_i^{(-k)} = \begin{bmatrix} x_{i,-k+1} \\ \vdots \\ x_{i,i} \\ \vdots \\ x_{i,T-k} \end{bmatrix}_{(T\times 1)} \qquad \epsilon_i = \begin{bmatrix} \epsilon_{i,1} \\ \vdots \\ \epsilon_{i,i} \\ \vdots \\ \epsilon_{i,T} \end{bmatrix}_{(T\times 1)}$$

the residual sum of squares is athen defined as:

$$RSS_0 = y'Qy - (y'QZ)(Z'QZ)^{-1}(Z'Qy) \tag{3.19}$$

where the matrix $[\tilde{W} : X]$ the vector y and the Within operator Q are respectively de fined by:

$$\tilde{W} = \begin{bmatrix} W_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & & W_i & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & 0 & \dots & W_N \end{bmatrix}_{NT\times NK} \qquad X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{(NT\times K)} \qquad , y = \begin{bmatrix} y_1^{(0)} \\ \vdots \\ y_i^{(0)} \\ \vdots \\ y_N^{(0)} \end{bmatrix}_{(NT\times 1)} \qquad ,$$

$Q = I_N \otimes Q_T = [I_N \otimes (I_T - \frac{1}{T} ee')]_{NT\times NT}$
where e denotes a $(T, 1)$ unit vector and where W_i and X_i are respectively defined by:

$$W_i = \left[y_i^{(1)} : \quad \dots \quad : y_i^{(i)} : \quad \dots \quad : y_i^{(K)} \right]_{T\times K} \qquad X_i = \left[x_i^{(1)} : \quad \dots \quad : x_i^{(i)} : \quad \dots \quad : x_i^{(K)} \right]_{T\times K}$$

The value of RSS_1 get under the alternative $\beta_i^{(k)} \neq \beta_i^{(k)}$, is computed as the sum of the residual sum of squares of individual estimations:

$$RSS_1 = \sum_{i=1}^{N} RSS_{1,i} \tag{3.20}$$

Let us denote $Z_i = [W_i : X_i], Z_{c,i} = [e : W_i : X_i]$ and $y_i = y_i^{(0)}$, the residual sum of squares of individual estimations is then defined as:

$$RSS_{1,i} = y_i' Q_T y_i - y_i' Q_T Z_i (Z_i' Q_T Z_i)^{-1} Z_i' Q_T y_i = y_i' y_i - y_i' Z_{ci} (Z_{ci}' Z_{ci})^{-1} Z_{ci}' y_i \tag{3.21}$$

Under assumptions **A1**, and particularly the normality of $\epsilon_{i,t}$, the F_H statistic has a Fischer distribution with $K(N-1)$ and $N(T-2K-1)$ degrees of freedom. If the realization of this statistic is not significant, the homogeneity hypothesis is accepted: the parameters $\beta_i^{(k)}$ can be restricted to be common for all units, as in *Holtz-Eakin and all. (1988)*.

In model 3.7, the homogeneity test on parameters $\beta_i^{(k)}$ is built with individual values of $\gamma_i^{(k)}$ under the null and the alternative hypothesis. Such an assumption allows us to distinguish between the homogeneity of $\beta_i^{(k)}$, which is central in the Granger causality test, and the homogeneity of the autoregressive parameters of $y_{i,t}$. However in a second step, it could be also possible to test jointly the homogeneity of all parameters as in a standard specification test (Hsiao 1986). If we assume that autoregressive parameters $\gamma_i^{(k)}$ are homogenous across groups under the null H_0, the RSS_0 is also defined by equation 3.19 but with $Z = [W : X]$. Under the alternative H_1, the value of RSS_1 is then defined by:

$$RSS_1 = y'Qy - (y'Q\tilde{Z})(\tilde{Z}'Q\tilde{Z})^{-1}(\tilde{Z}'Qy) \tag{3.22}$$

with $\tilde{Z} = [W : \tilde{X}]$, where $\tilde{X}$ and W are respectively defined by:

$$\tilde{X} = \begin{bmatrix} X_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & & X_i & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & 0 & \dots & X_N \end{bmatrix}_{NT \times NK} \qquad W = \begin{bmatrix} W_1 \\ \vdots \\ W_i \\ \vdots \\ W_N \end{bmatrix}_{(NT \times p)}$$

In this case, under assumptions **A1**, the F_H statistic has a Fischer distribution with $K(N-1)$ and $NT - N(K+1) - K$ degrees of freedom.

3.4.2 HNC hypothesis test

The second step of our procedure consists in *homogenous non causality* hypothesis *(HNC)*. We test whether or not the regression slope coefficients associated to $x_{i,t-k}$ are null for all individual i and all lag k. However, as in the unit root tests literature, the alternative hypothesis would be different according the homogeneity of parameters $\beta_i^{(k)}$ and so the results of the homogeneity test.

3.4.2.1 HNC hypothesis test under homogeneity

First, we assume that the regression slope coefficients associated to $x_{i,t-k}$ are homogenous for all units :$\beta_i^{(k)} = \beta^{(k)} =, k = 1, .., K$. In this case, the *homogenous non causality hypothesis (HNC)* test is identical to those proposed by *Holtz-Eakin, Newey and Rosen (1988)*. Under the alternative hypothesis, we assume the homogeneity of parameters $\beta_i^{(k)}$. That is why this HNC test is very similar to the unit root test proposed by *Levin and Lin (1992)*.
$H_0 : \beta^{(k)} = 0, \quad \forall k = 1, ..., K$
$H_1 : \beta^{(k)} \neq 0, \quad \exists k = 1, ..., K$
In order to test these K linear restrictions, we compute the following F^a_{HNC} statistic:

$$F^a_{HNC} = \frac{(RSS_2 - RSS_0)/K}{RSS_0/[NT - N(K+1) - K]} \tag{3.23}$$

where RSS_0 has been previously defined (equation 3.19) and RSS_2 denotes the restricted sum of squared residual obtained under H_0 of the model 3.7. If we consider heterogenous autoregressive parameter $\gamma_i^{(k)}$, RSS_2 is then defined as the sum of individual sum of squares get under H_0 :

$$RSS_2 = \sum_{i=1}^{N} RSS_{2,i} \tag{3.24}$$

Let us denote $W_{ci} = [e : W_i]$, the sum of squares get under H_0 for the i^{th} cross section unit is:

$$RSS_{2,i} = y_i' Q_T y_i - y_i' Q_T W_i (W_i' Q_T W_i)^{-1} W_i' Q_T y_i \tag{3.25}$$
$$= y_i' y_i - y_i' W_{ci} (W_{ci}' W_{ci})^{-1} W_{ci}' y_i \tag{3.26}$$

Under assumptions **A1**, the F^a_{HNC} statistic has a Fischer distribution with K and $NT - N(K+1) - K$ degrees of freedom. If the realization of this statistic is not signi ficantly different from zero, the Homogeneity Non Causality hypothesis is accepted : the variable x does not cause y for all the individuals i. There does not exist any individual causality relationships from x to y.

As for the homogeneity test, we can also assume that autoregressive parameters $\gamma_i^{(k)}$ are homogenous under the null and the alternative hypothesis. In this case, the RSS_0 is also defined by equation 3.19 but with $Z = [W : X]$, and the residual sum of square under the null HNC is equal to $RSS_2 = y'Qy - (y'QW)(W'QW)^{-1}(WQy)$. In this case, the F^a_{HNC} statistic has a Fischer distribution with K and $NT - N - 2K$ degrees of freedom under the HNC null hypothesis.

3.4.3 HNC hypothesis test under heterogeneity

Now, we assume that the conclusion of the homogeneity test is the rejection of the null hypothesis: the regression slope coefficients $\beta_i^{(k)}$ associated to $x_{i,t-k}$ are heterogenous. The null hypothesis of the Homogenous Non Causality (HNC) test still unchanged:

$$H_0 : \beta_i^{(k)} = 0 \qquad \forall i = k, ..., K, \quad \forall i = 1, ..., N \tag{3.27}$$

However, the alternative is now given by:

$$H_1 : \beta_i^{(k)} = 0 \qquad \forall k = 1, ..., K, \quad \forall i = 1, ..., N_1 \tag{3.28}$$
$$H_1 : \beta_i^{(k)} \neq 0 \qquad \exists k = 1, ..., K, \quad \forall i = N_1 + 1, N_1 + 2..., N \tag{3.29}$$

Then, in this case under the alternative there exists a subgroup of units (with dimension N_1) for which the variable x does not Granger cause the variable y and an another subgroup (dimension $N - N_1$) for which the variable x Granger causes y, since at least one regression slope coefficient associated to $x_{i,t-k}$ is different from zero. The structure of this test is similar to the unit root test of *Im, Pesaran and Shin (1997)*. That is why, we can also use in this context a statistic of test based on the average of individual F statistics associated to the test of the non causality hypothesis for units $i = 1, .., N$. For simplicity, we propose here to use the average of individual Wald statistics rather than average of F statistics. The statistic W^b_{HNC} associated to the null hypothesis of the Homogenous Non Causality (HNC) under heterogeneity is then defined as:

$$W^b_{HNC} = \frac{K}{N} \sum_{i=1}^{N} F_i^b \tag{3.30}$$

where F_i^b denotes the individual Fischer statistics for the ith cross section unit, associated to the test $H_0 : \beta_i^k = 0, \forall k = 1, .., K$.

$$F_i^b = \frac{(RSS_{2,i} - RSS_{1,i})/K}{RSS_{1,i}/(T - 2K - 1)} \quad \forall i = 1, ..., N \tag{3.31}$$

where $RSS_{1,i}$,i and $RSS_{2,i}$, are define as above in sections 3.4.1 and 3.4.2.1 respectively.

The challenge is then to determine the distribution of the W^b_{HNC} statistic. Let us denote $Z_{ci} = [e : W_i : X_i]$ and $\theta_i = (\alpha_i, \gamma_i^{(1)}, ..., \gamma_i^{(k)}, \beta_i^{(1)}, ..., \beta_i^{(k)})'$ the vector of parameters of model 3.7. Th e HNC hypothesis can be expressed as $R\theta_i = 0$ where R is a $(K, 2K+1)$ matrix with $R = [0 : I_K]$. The product KF_i^b will also be defined as (cf. appendix A.1):

$$W_i^b = KF_i^b = (\frac{\tilde{\epsilon}_i' \Phi_i \tilde{\epsilon}_i}{\tilde{\epsilon}_i' M_i \tilde{\epsilon}_i})(T - 2K - 1) \tag{3.32}$$

where $\tilde{\epsilon}_i = \frac{\epsilon_i}{\sigma_{\epsilon,i}}$ and where the matrix Φ_i and M_i are semi-positive definite, symmetric and idempotent $(T \times T)$ matrix.
$\Phi_i = Z_{ci}(Z_{ci}'Z_{ci})^{-1}R'[R(Z_{ci}'Z_{ci})^{-1}R']^{-1}R(Z_{ci}'Z_{ci})Z_{ci}'$ and $M_i = I_T - Z_{ci}(Z_{ci}'Z_{ci})^{-1}Z_{ci}'$

First, let us consider the asymptotic distribution of the Wald statistic W_i^b . In a non dynamic model, the normality assumption **A1** would be sufficient to establish the fact for all T, the Wald statistic has a chi-squared distribution with K degrees of freedom. But in a dynamic model, this result can only be achieved asymptotically *(Hamilton 1994)*. Given that under **A1** the OLS estimate $\hat{\theta}_i$ is convergent. Under assumption **A1**, the vector $\tilde{\epsilon}_i$ is distributed across a N (0, 1) . Since Φ_i is idempotent, the quadratic form $\tilde{\epsilon}_i' \Phi_i \tilde{\epsilon}_i$ is distributed as chi- squared with a number of degrees of freedom equal to the rank of Phi_i. The rank of Phi_i is equal to its trace, that is to say K (cf. appendix A.1). With or without the normality assumption in **A1**, we have here fore each individual:

$$W_i^b \xrightarrow[T \rightarrow \infty]{L} \mathcal{X}^2(K) \tag{3.33}$$

Asymptotically, individual Wald statistics W_i^b converge toward an identical chi-squared distribution with finite second order moments, $\forall i = 1, .., N$. However, this convergence result can not be achieved for any time dimension T, even if we assume the normality of residuals. The issue is then to show that for all T, the individual Wald statistics have finite second order moments. Let us consider the expression 3.32 this is a ratio of two quadratic forms in a standard normal vector under assumption. As in *Im, Pesaran and Shin (1997)*, we can show that W_i^b has finite second order moment for all T, without any convergence in distribution results, from the *Magnus (1990) theorem*. This theorem allows us to establish that the s^{th} moment of the ratio $E[\frac{(\tilde{\epsilon}_i' \Phi_i \tilde{\epsilon}_i)}{(\tilde{\epsilon}_i' M_i \tilde{\epsilon}_i)}]^s$ exists as soon as $0 \leq s \leq rank(M_i)/2$. In our context, we have:
$rank(M_i) = trace[I_T - Z_{ci}(Z_{ci}'Z_{ci})^{-1}Z_{ci}'] = trace(I_T) - trace[(Z_{ci}'Z_{ci})^{-1}Z_{ci}'Z_{ci}] = trace(I_T) - trace(I_{2K+1}) = T - 2K - 1$
This condition is also satis fied for W_i^b :

$$E[(W_i^b)] = E[\frac{(\tilde{\epsilon}_i' \Phi_i \tilde{\epsilon}_i)}{(\tilde{\epsilon}_i' M_i \tilde{\epsilon}_i)}]^s (T - 2K - 1), \quad \text{exists if}, s \leq \frac{T - 2K - 1}{2} \tag{3.34}$$

So, for all cross unit $i = 1, .., N$, for all time dimension T, the condition of the *Magnus* *theorem (1990)* that establishes the fact that the second order moments ($s = 2$) of the Wald statistic W_i^b are finite is:

$$T \geq 5 + 2K \tag{3.35}$$

Hence, individual Wald statistics KF_i^b are identically distributed with finite second order moments. Before applying a standard central limit theorem, it can be proved that these individual statistics are independent. For that, let us consider the expression 3.32. Under assumption **A2**, residual ϵ_i and ϵ_j for$i \neq j$ are independent at all date. Then, under assumption **A3**, the Wold decompositions associated to the vectorial processes Z_{ci} and Z_{cj} are also independent. Then, Wald statisticsKF_i^b and KF_j^b are independently distributed.

3.4.4 HENC hypothesis test

Let us assume that the regression slope coefficients associated to $x_{i,t-k}$ are heterogenous and the HNC is rejected. In this case, the last step of the procedure consists in determining the size N_1 of the subgroup of units to which there is no Granger causality from x to y. If $N_1 = 0$, we get the HEterogenous Causality (HEC) case. If $N_1 > 0$, the variable x Granger causes y for all units, but the dynamic relations between both variables is heterogenous: this is the HEterogenous Non Causality ($HENC$) case.

Chapter 4

DATA ANALYSIS AND MODELS

4.1 Introduction

Since the early 1960s scholars and policy makers alike, have all shown great interest in the possible relationship between exports and economic growth. The motivation is clear. Should a country promote exports to speed up economic growth or should it primarily focus on economic growth, which in turn will generate exports? There are basically four propositions. According to the export - led growth (ELG) hypothesis, export activity leads economic growth. Trade theory provides several plausible explanations in favour of this idea. For example, export promotion directly encourages the production of goods for exports. This may lead to further specialisation in order to exploit economies of scale and the nation's comparative advantages. Moreover, increased exports may permit the imports of high quality products and technologies, which in turn may have a positive impact on technological change, labour productivity, capital efficiency and, eventually, on the nation's production. The second proposition, the growth - driven exports (GDE) hypothesis, postulates a reverse relationship. It is based on the idea that economic growth itself induces trade flows. It can also create comparative advantages in certain areas leading to further specialisation and facilitating exports. These two propositions do not exclude each other, so the third notion is a feedback relationship between exports and economic growth. Finally, it is also possible, though unlikely, that there is no relationship or just a simple contemporaneous, maybe spurious relationship, between these two variables. Also the developments of world economies puts the financial systems as leaders in the world exchange market. In the open economies, foreign exchange rate policies are one of the most important macroeconomic indicators, because of the fact that they affect the business world's investment decisions.

Table 4.1: List of the 27 develoving African countries considered

Algeria	*Benin*	**Botswana**	*Burkina Faso*	*Burundi*	**Cameroon**
Central Africa Rep	*Congo, Rep*	**Cote d'Ivoire**	**Egypt**	**Gabon**	**Ghana**
Kenya	**Lesotho**	*Madagascar*	*Malawi*	**Mauritania**	**Morocco**
Niger	**Nigeria**	*Rwanda*	*Senegal*	**South Africa**	**Sudan**
Swaziland	*Togo*	*Uganda*			

Notes: Using the 2014 World Bank classification, low- income countries are italicized, while the middle-income countries are indicated in bold font.

4.2 Data:

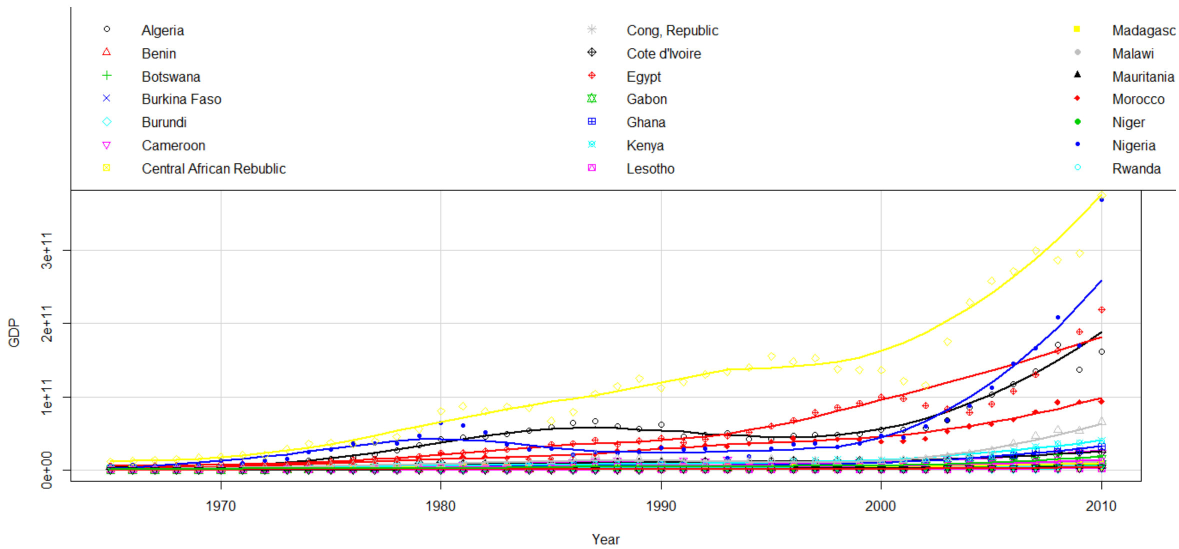
Based on the $World\ Bank's$ 2013 definition: (Developing countries are defined according to their Gross National Income (GNI) per capita per year. Countries with a GNI of US$ 11,905 and less are defined as developing (specified by the World Bank, 2013).)and International Monetary Fund,IMF, 2014 classification, there are about 139 developing countries, of which 51 of them are from Africa. The main purpose is paper is to test for Granger causality between the logarithms of Official exchange rate (LCU, Lcal Currency Unit) per US$, period average), Exports of goods and services (% of GDP) and GDP (current US$) in 27 developing countries in Africa from 1965 to 2010 inclusive. These 27 countries were chosen only due to data limitations. These countries form a balanced panel, each with the trivariate variables from 1965 through 2010. Lists of countries are provided in Table 4.1.

4.2.1 Gross Domestic Product (GDP)

The gross domestic product (GDP) is one of the primary indicators used to gauge the health of a country's economy. It represents the total dollar value of all goods and services produced over a specific time period; you can think of it as the size of the economy. Usually, GDP is expressed as a comparison to the previous quarter or year. For example, if the year-to-year GDP is up 3%, this is thought to mean that the economy has grown by 3% over the last year. Measuring GDP is complicated (which is why we leave it to the economists), but at its most basic, the calculation can be done in one of two ways: either by adding up what everyone earned in a year (income approach), or by adding up what everyone spent (expenditure method). Logically, both measures should arrive at roughly the same total.

The income approach, which is sometimes referred to as GDP(I), is calculated by adding up total compensation to employees, gross profits for incorporated and non incorporated firms, and taxes less any subsidies. The expenditure method is the more common approach and is calculated by adding total consumption, investment, government spending and net exports. As one can imagine, economic production and growth, what GDP represents, has a large impact on nearly everyone within that economy. For example, when the economy is healthy, you will typically see low unemployment and wage increases as businesses demand labor to meet the growing economy. A significant change in GDP, whether up or down, usually has a significant effect on the stock market. It's not hard to understand why: a bad economy usually means lower profits for companies, which in turn means lower stock prices. Investors really worry about negative GDP growth, which is one of the factors economists use to determine whether an economy is in a recession. Figure 4.1 shows the plot of the gdp for the 27 countries under consideration, whilst the country - wise plot is depicted in figure 4.2

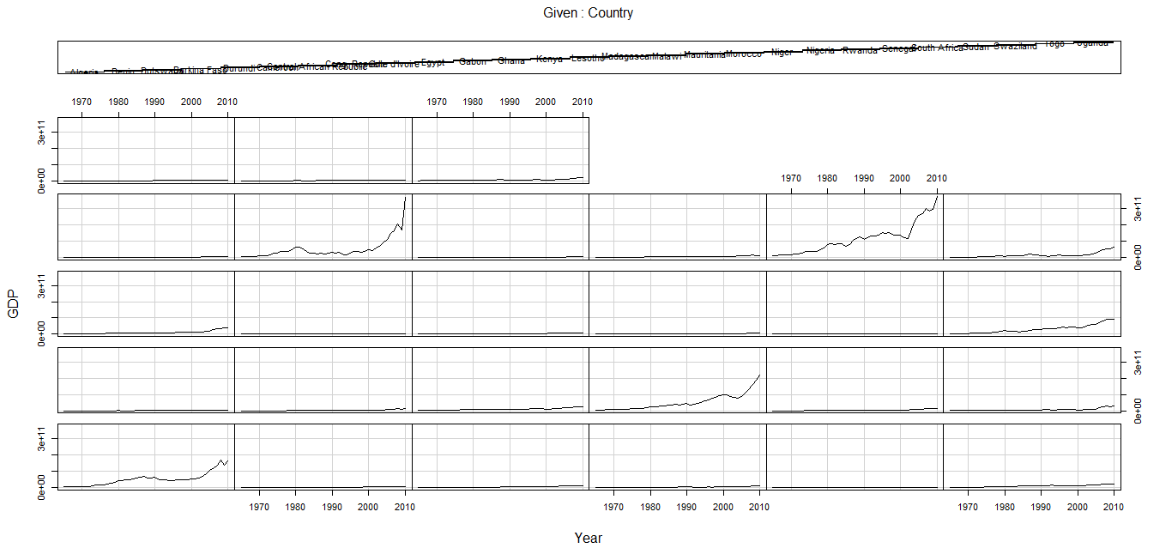
Figure 4.1: Joint GDP plot of the 27 countries all together



4.2.1.1 Real GDP

One thing people want to know about an economy is whether its total output of goods and services is growing or shrinking. But because GDP is collected at current, or nominal, prices, one cannot compare two periods without making adjustments for inflation. To determine "real" GDP, its nominal value must be adjusted to take into account price changes to allow us to see whether the value of output has gone up because more is being produced or simply because prices have increased. A statistical tool called the price deflator is used to adjust GDP from nominal to constant prices. GDP is important because it gives information about the size of the economy and how an economy is performing. The growth rate of real GDP is often used as an indicator of the general health of the economy. In broad terms, an increase in real GDP is interpreted as

Figure 4.2: Individual GDP plots of the countries under consideration



a sign that the economy is doing well. When real GDP is growing strongly, employment is likely to be increasing as companies hire more workers for their factories and people have more money in their pockets. When GDP is shrinking, as it did in many countries during the recent global economic crisis, employment often declines. In some cases, GDP may be growing, but not fast enough to create a sufficient number of jobs for those seeking them. But real GDP growth does move in cycles over time. Economies are sometimes in periods of boom, and sometimes in periods of slow growth or even recession (with the latter often defined as two consecutive quarters during which output declines). In the United States, for example, there were six recessions of varying length and severity between 1950 and 2011. The National Bureau of Economic Research makes the call on the dates of U.S. business cycles.

4.2.1.2 Comparing GDPs of two countries

GDP is measured in the currency of the country in question. That requires adjustment when trying to compare the value of output in two countries using different currencies. The usual method is to convert the value of GDP of each country into U.S. dollars and then compare them. Conversion to dollars can be done either using market exchange rates - those that prevail in the foreign exchange market - or purchasing power parity (PPP) exchange rates. The PPP exchange rate is the rate at which the currency of one country would have to be converted into that of another to purchase the same amount of goods and services in each country. There is a large gap between market and PPP - based exchange rates in emerging market and developing countries. For most emerging market and developing countries, the ratio of the market and PPP U.S. dollar exchange rates is between 2 and 4. This is because nontraded goods and services tend to be cheaper in low - income than in high - income countries, for example, a haircut in New York is more expensive than in Bishkek - even when the cost of making tradable goods, such as machinery, across two countries is the same. For advanced economies, market and PPP exchange rates tend to be much closer. These differences mean that emerging market and developing countries have a higher estimated dollar GDP when the PPP exchange rate is used. The *IMF* publishes an array of GDP data on its website http://www.imf.org. International institutions such as the IMF also calculate global and regional real GDP growth. These give an idea of how quickly or slowly the world economy or the economies in a particular region of the world are growing. The aggregates are constructed as weighted averages of the GDP in individual countries, with weights reflecting each country's share of GDP in the group (with PPP exchange rates used to determine the appropriate weights).

4.2.1.3 What GDP does not reveal

It is also important to understand what GDP cannot tell us. GDP is not a measure of the overall standard of living or well-being of a country. Although changes in the output of goods and services per person (GDP per capita) are often used as a measure of whether the average citizen in a country is better or worse off, it does not capture things that may be deemed important to general well-being. So, for example, increased output may come at the cost of environmental damage or other external costs such as noise. Or it might involve the reduction of leisure time or the depletion of nonrenewable natural resources. The quality of life may also depend on the distribution of GDP among the residents of a country, not just the overall level. To try to account for such factors, the United Nations computes a Human Development Index, which ranks countries not only based on GDP per capita, but on other factors, such as life expectancy, literacy, and school enrollment. Other attempts have been made to account for some of the shortcomings of GDP, such as the Genuine Progress Indicator and the Gross National Happiness Index, but these too have their critics.

4.2.1.4 Limitations of GDP Statistics

Basically, GDP growth measures the output of an actual economy. However, to gain a better understanding of average living standards we need to look at the growth of GDP per capita. For example, if one country has GDP growth of 4%, but the population increases in size by 4%, then the average citizen will have the same income. Another country, could have zero GDP growth but, if the population is declining then the average citizen will be better off. The economist points out that when we use GDP per capita growth figures, Japan has actually outperformed the US economy. Because although actual GDP has increased faster in the US, this has been boosted by a growing population. It means that other fast growing economies like India, should also have their GDP statistics treated with caution. Although, growth rates are much higher in India, it is partly boosted by a rapidly rising population.

4.2.2 Exports of goods and services in % of GDP

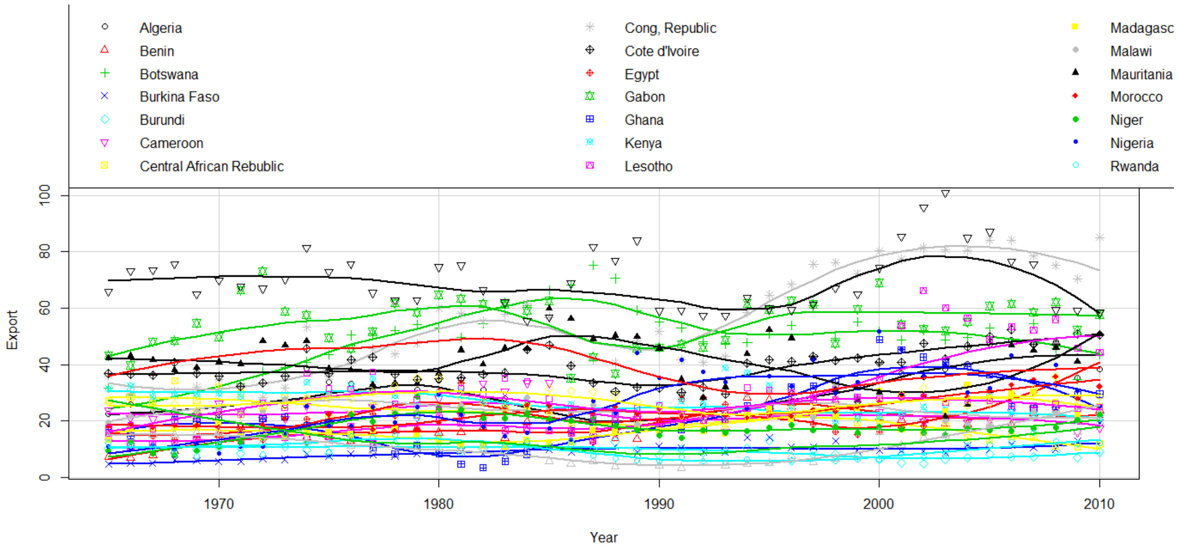
This indicator is the value of exports of goods and services divided by the GDP in current prices. Exports of goods and services (% of gdp) of a given country, according to the world bank represent the value of all goods and other market services provided to the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude compensation of employees and investment income (formerly called factor services) and transfer payments.

The term export means shipping in the goods and services out of the jurisdiction of a country. The seller of such goods and services is referred to as an “exporter" and is based in the country of export whereas the overseas based buyer is referred to as an "importer". In international trade, “exports" refers to selling goods and services produced in the home country to other markets. Export of commercial quantities of goods normally requires involvement of the customs authorities in both the country of export and the country of import. The advent of small trades over the internet such as through *Amazon* and *eBay* have largely bypassed the involvement of Customs in many countries because of the low individual values of these trades. Nonetheless, these small exports are still subject to legal restrictions applied by the country of export. An export's counterpart is an import.
Trade in goods and services is defined as change in ownership of material resources and services between one economy and another. The indicator comprises sales of goods and services as well as barter transactions or goods exchanged as part of gifts or grants between residents and non-residents. It is measured in million USD and percentage of GDP for net trade and also annual growth for exports and imports. Figure 4.3 shows the plot of the export of goods and services for the 27 countries under consideration, whilst the country - wise plot is depicted in figure 4.4

4.2.2.1 BREAKING DOWN ‘Export’

Exports are one of the oldest forms of economic transfer and occur on a large scale between nations that have fewer restrictions on trade, such as tariffs or subsidies. Most of the largest companies operating in advanced economies derive a substantial portion of their annual revenues from exports to other countries. The ability to export goods helps an economy to grow, by selling more overall goods and services. One of the core

Figure 4.3: Joint Exports of goods and services in % of GDP plot of the 27 countries all together

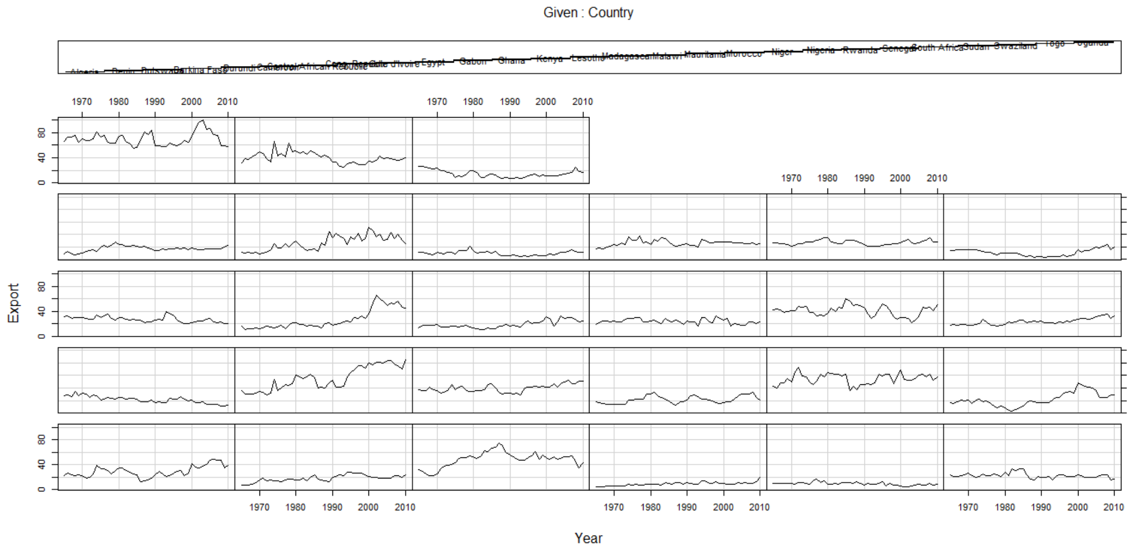


functions of diplomacy and foreign policy within governments is to foster economic trade in ways that benefit both parties involved. Exports are a crucial component of a country's economy. Not only do exports facilitate international trade, they also stimulate domestic economic activity by creating employment, production and revenues. As of 2014, the world's largest exporting countries in terms of dollars are China, the United States, Germany, Japan and the Netherlands. China has exports of approximately $2.3 trillion, primarily exporting electronic equipment and machinery. The United States exports approximately $1.6 trillion, primarily exporting capital goods. Germany has exports of approximately $1.5 trillion, primarily exporting motor vehicles. Japan has exports of approximately $684 billion, primarily exporting motor vehicles. Finally, the Netherlands has exports of approximately $672 billion, primarily exporting machinery and chemicals.

4.2.3 Exchange Rate

In finance, an exchange rate between two currencies is the rate at which one currency will be exchanged for another. It is also regarded as the value of one country's currency in terms of another currency. Exchange rates are determined in the foreign exchange market, which is open to a wide range of different types of buyers and sellers, and where currency trading is continuous. The *spot exchange rate* refers to the current exchange rate. The *forward exchange rate* refers to an exchange rate that is quoted and traded today but for delivery and payment on a specific future date. In the retail currency exchange market, different buying and selling rates will be quoted by money dealers. Most trades are to or from the local currency. The buying rate is the rate at which

Figure 4.4: Individual Exports of goods and services in % of GDP plots of the countries under consideration



money dealers will buy foreign currency, and the selling rate is the rate at which they will sell that currency. The quoted rates will incorporate an allowance for a dealer's margin (or profit) in trading, or else the margin may be recovered in the form of a commission or in some other way. Different rates may also be quoted for cash (usually notes only), a documentary form (such as traveler's cheques) or electronically (such as a credit cardpurchase). The higher rate on documentary transactions has been justified as compensating for the additional time and cost of clearing the document. On the other hand, cash is available for resale immediately, but brings security, storage, and transportation costs, and the cost of tying up capital in a stock of banknotes (bills).

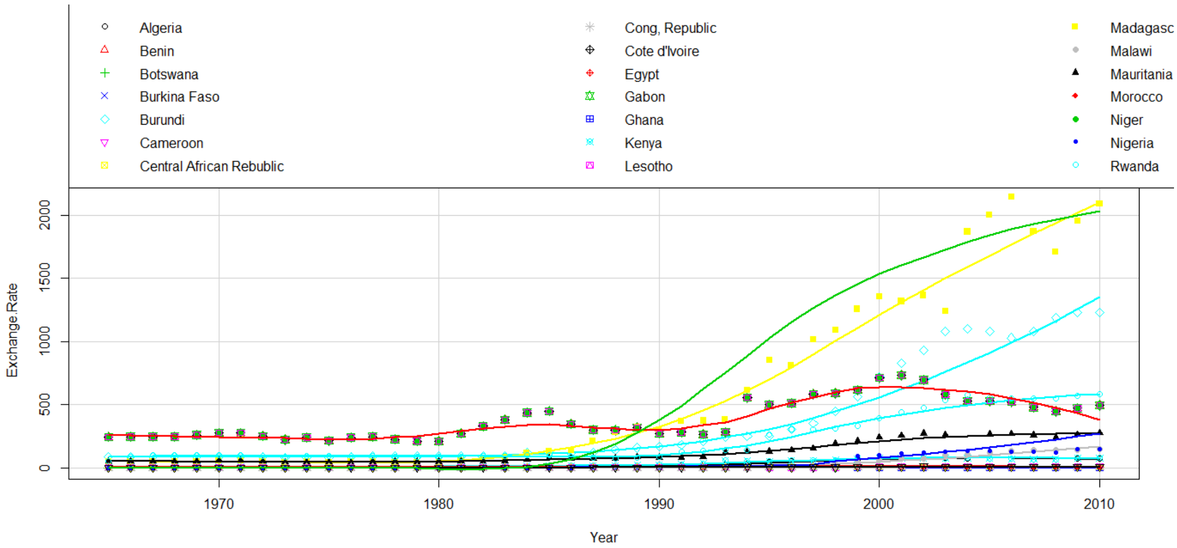
4.2.3.1 Official exchange rate (LCU per US$, period average)

Official exchange rate refers to the exchange rate determined by national authorities or to the rate determined in the legally sanctioned exchange market. It is calculated as an annual average based on monthly averages (local currency units relative to the U.S. dollar). Figure 4.5 shows the plot of the export of goods and services for the 27 countries under consideration, whilst the country - wise plot is depicted in figure 4.6

4.2.3.2 Development Relevance:

In a market - based economy, household, producer, and government choices about resource allocation are influenced by relative prices, including the real exchange rate, real wages, real interest rates, and other prices in the economy. Relative prices also largely reflect these agents' choices. Thus relative prices convey vital information about the interaction of economic agents in an economy and with the rest of the world.

Figure 4.5: Joint Official exchange rate (LCU per US$, period average) plot of the 27 countries all together



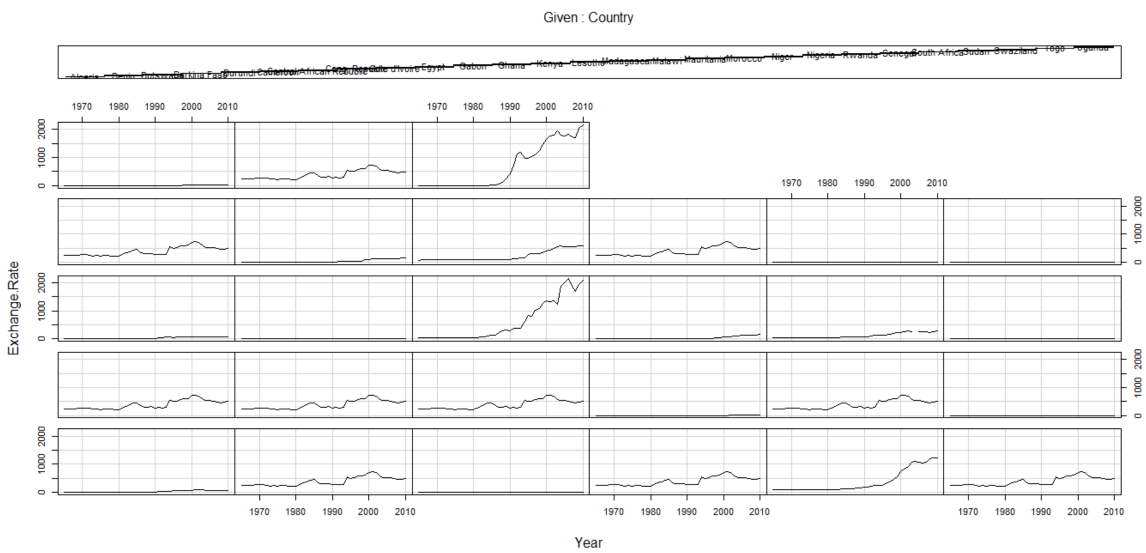
4.2.3.3 Limitations and Exceptions:

Official or market exchange rates are often used to convert economic statistics in local currencies to a common currency in order to make comparisons across countries. Since market rates reflect at best the relative prices of tradable goods, the volume of goods and services that a U.S. dollar buys in the United States may not correspond to what a U.S. dollar converted to another country's currency at the official exchange rate would buy in that country, particularly when nontradable goods and services account for a significant share of a country's output. An alternative exchange rate - the purchasing power parity (PPP) conversion factor - is preferred because it reflects differences in price levels for both tradable and nontradable goods and services and therefore provides a more meaningful comparison of real output.

4.2.3.4 Statistical Concept and Methodology:

The exchange rate is the price of one currency in terms of another. Official exchange rates and exchange rate arrangements are established by governments. Other exchange rates recognized by governments include market rates, which are determined largely by legal market forces, and for countries with multiple exchange arrangements, principal rates, secondary rates, and tertiary rates. The country with the highest value in the world is Zimbabwe, with a value of 6,723,052,000.00. The country with the lowest value in the world is Kuwait, with a value of 0.30.

Figure 4.6: Individual Official exchange rate (LCU per US$, period average) plots of the countries under consideration



4.3 Modelling

In order to investigate the export - growth - exchange rate nexus, we consider a panel with 27 cross - sections (countries) and 45 years, by following the VAR method and VECM methods. Also due to contemporaneous correlations across countries, some additional panel information were obtained (i.e. the equations represent a SUR (Seemingly Unrelated Related) system). The variables considered are: GDP (current US $), exports of goods and services in % GDP and exchange rate. The natural log og these variables is used and is denoted by LGDP, LEXP and LEXC respectively. These transformations was done because, coefficients can be understood as elasticities of a Cobb-Douglas function. This function which is probably the most common one used among economists to analyze issues regarding microeconomic behaviour (consumerspreferences, technology, production functions) and macroeconomic issues (economic growth). The elasticity term is used to describe the degree of response of a change of a variable with respect to another. The summary statistics can be found on appendix C .2.

4.3.1 Panel Unit Root Testing

We begin our analysis by performing descriptive statitics on our data set. The results of this analysis is the appendix. Next we analyze the statistical properties of our data. We test whether our panel data has unit root or not. Following the test procedure described in the previuos chapter, section 2.2.1 and 2.2.2, the test is performed and summarized in Table 4.2

The test was considered for two cases, when there was no intercept and trend and

Table 4.2: Panel Unit root test

	Ch-square distribution				
Test Procedure	Variable	Levels Statistics	p value	First Diff static	p value
Levin, Lin & Chu	LEXP	0.60500	0.7274	-37.1358	0.0000
	LGDP	17.3932	1.0000	-21.7771	0.0000
	LEXC	5.17084	1.0000	-22.7665	0.0000
ADF - Fisher Chi-square	LEXP	23.8505	0.9999	1891.68	0.0000
	LGDP	0.14131	1.0000	563.317	0.0000
	LEXC	18.2106	1.0000	616.461	0.0000
PP - Fisher Chi-square	LEXP	20.5096	1.0000	2500.56	0.0000
	LGDP	0.09318	1.0000	582.532	0.0000
	LEXC	18.9361	1.0000	615.639	0.0000

Notes: Lag selection was based on the Schwarz Information Criterion, Newey-West automatic bandwidth selection and Bartlett kernel.

secondly when there was an intercept and trend in the model. Since the lag selection is very essential step in the modelling processes, the Schwarz Information Criterion, together with Newey-West automatic bandwidth selection and Bartlett kernel were used to obtain the optimal lag to include in the model . As can be observed from table 4.2, all the three test procedures considered fail to reject the null hypothesis of unit root for all the variable at their levels. The first difference of all the three variables are significant at 5% which means that the variables are stationary at this instance. Evident from the Levin, Lin & Chu , ADF and PP test imply that all the variables are integrated of order one, $I(1)$ with their difference being stationary.

4.3.2 Panel Cointegration Testing

Since all the variables are integrated of the same order $I(1)$, we perform the cointrgration test using the Fisher - Johansen test statistic for a possible cointegration relation(s) among the variables. Table 4.3 summarize the results of the test for cointegration between the three variables for all the panel data set based on Cointegration Rank Test (Trace and Maximum Eigenvalue).The trace test and maximum eigenvalue test described in section 5.1 with lag order 2 are used. The test is performed unrestricted constant which allow for the drift vector Γ in the model.

From table 4.3, it can be seen that, the result indicate cointegration relationship among the variables. The first null hypothesis of no cointegrated equations can be

Table 4.3: Johansen Fisher Panel Cointegration Test

Hypothesized No. of CE(s)	Fisher Stat*. (from trace test)	p value	Fisher Stat*. (from max-eigen test)	p value
None	100.1	0.0001	93.09	0.0008
At most 1	43.64	0.8421	41.75	0.0008
At most 2	30.04	0.9966	30.04	0.9966

Notes: *Probabilities are computed using asymptotic Chi-square distribution..

rejected since the p - value is significant for both the trace and maximum eigen test. The next two null hypotheses states that there is at least one cointegrated equations. We cannot reject these hypotheses since the test statistic is not significant evident from table 4.3. This means that there exist 1 and 2 cointegration relations among the GDP, exchange rate and export variables. These test results suggest to us to study the long term relationship that exist among the variables and the direction of causality that might run among the variables in the long term. We estimate the VECM model which is consistent with the foregoing analysis and discuss the long term relationship among the variables.

4.3.3 Vector Error Correction Model (VECM)

A very essential part of this thesis is the estimation of the correct model that best describe the Data Generating Process of our variables to ensure a good statistical inference. Based on the Fisher Johansen test results, we estimate a VECM which provides means to analyze the dynamic disequilibrium of our variables both in the short run and the long run. The results will be presented as follows:

- First we present the cointegration equation for the overall panel data in the following ways:
  - building the long run analysis cointgration equation of our model.
  - then we build the short term dynamic disequilibrium of the variables in the VECM
- Secondly, we present the cointegration equation for the country - by - country in the following ways:

- building the long run analysis cointgration equation of our country wise model.
- then we build the short term dynamic disequilibrium of the variables in the VECM country wise

Table 4.4: Vector Error Correction Estimates

Dependent Variable	Cointegrating equation:	Estimated value	t - statistic
LEXPORT	LEXP	1.0000	- - - -
	LGDP	-0.113969	-1.03982
	LEXC	0.326769	5.98545
	CONSTANT	-1.708809	- - - -
GDP	LGDP	1.000000	- - - -
	LEXP	-8.774304	-3.38421
	LEXC	-2.867172	-6.00535
	CONSTANT	14.99361	- - - -
EXCHANGE RATE	LEXC	1.0000	- - - -
	LGDP	-0.348776	-1.04383
	LEXP	3.060265	3.38602
	CONSTANT	-5.229408	- - - -

Notes: Tests performed after adjustments using the AIC and the SIC

The long term equation of the VEC model can be estimated from table 4.4 as :

$$LEXP_{t-1} = 1.708809 + 0.113969 LGDP_{t-1} - 0.326769 LEXC_{t-1} \tag{4.1}$$

Equation 4.1 shows a positive correlation between export of goods and services and and gross domestic product and a negative correlation between exchange rate and export. As per equation 4.1, a 1% increase in GDP leads to 11% rise in the export of good and services, while a 1% increase in exchange rate leads to a decrease in export by 32.7% in the long run.
The long run model for GDP as dependent variable is given by:

$$LGDP_{t-1} = -14.99361 + 8.774304 LEXP_{t-1} + 2.867172 LEXC_{t-1} \tag{4.2}$$

Also equation 4.5 shows a positive correlation between export of goods and services and and gross domestic product as well as exchange rate and the gross domestic product.

As can be seen from equation 4.5, a 1% increase in the export of goods and services leads to 87.7% rise in the GDP, and a 1% increase in exchange rate leads to a rise in GDP by 28.7% in the long run. This means that an overall increase in export of goods and services will propel economic (GDP) growth more than that oil exchange rate Finally, the long run model for exchange rate dependent variable is given by:

$$LEXC_{t-1} = 5.229408 + 0.348776LGDP_{t-1} - 3.060265LEXP_{t-1} \tag{4.3}$$

The exchange rate equation 4.3 shows a positive correlation between GDP and exchange rate and a negative correlation between export and exchange rate.It can be seen from equation 4.3 that , a 1% increase in GDP will lead to a 34.9% rise in the Exchange rate, whilst a 1% increase in export leads to a decrease in exchange rate by 306% in the long run, while shows a very strong inverse corellation between exchange rate and export. This means that an overall increase in export of goods and services will propel economic (GDP) growth more than that oil exchange rate.

4.3.4 Causality Analysis

With the long run models developed, it is now time to investigate the granger causality between our variables and if possible determine the direction of the causality.

4.3.4.1 Long run causality

We try to investimate the direction of causality between our three variables of interest.

Table 4.5: Causality direction in the three models obtained

Dependent Variable in the models:	ECT (speed of adjustment towards long run equilm)	Estimated statistic	t - statistic	p value
EXPORT	$C(1)_{LEXP}$	-0.003248	-0.828247	0.40760
GDP	$C(1)_{LGDP}$	-0.000406	-1.051409	0.29330
EXCHANGE RATE	$C(1)_{LEXC}$	-0.009021	-1.51237	0.0001*

Notes: * means significant at 1%, 5% and 10% level of significance

From table 4.5 it can be seen that, when we consider the export variable, there is no granger causality from both GDP and exchange rate to export. That is, GDP and exchange rate does not jointly granger cause export. Similaarly, export and exchange rate does not jointly cause GDP. But there is a unidirectional causality from GDP and export to exchange rate, that is, GDP and export jointly cause exchange rate.

4.3.4.2 Short run causality

We then further demonstarte in Table 4.6, the coefficients of the short term dynamic disequilibrium of the variables in the VECM already an alyzed in section 4.3.3. Our main interest here is to determine the pairwise causal relationship and the direction between the three variables under consideration.Now focusing on the results in columns 2, 4 and 6, it can be seen from column 2 that both GDP and Exchange rate do not have any significant impart on export in the short run, it can also be seen that export has a negative significant effect on itself, that is a change in export produces a negative significant effect on itself for the first and second years. From column 4, the coefficient of export at the first and second lags are significant at 5% and 10% significance level. The positive values depict the positive correlation between the export and GDP in developing countries. We can also see that exchange rate has no significant effect on GDP. Also it is observed that GDP is only significant to itself in the first lag (year). Finally, from column 6, we can see that export is only significant to exchange rate in the first lag, the negative value here also depicts the negative correlation between the exchange rate and export, furthermore it can be seen that exchange rate negatively respond to DGP. There a positive effect of exchange rate to it self in the first and second years.

4.3.4.3 Pairwise causality

Since we are now equiped with all the information that we need, we try to investigate the pairwise causality between our three variables, that is we investigate the causality between GDP and Export, GDP and Exchange rate and Export and Exchange rate. The following table 4.7 summarizes our findings using the Wald test and we computed the chi - squared values instead of the F - statistics. As can be seen from table 4.7, there is no causality from either GDP or Exchange rate to Export. Moreover, we can see that there is a unidirectional causality from export to GDP. Finally, it is observed that, there is a one - way causality from GDP to exchange rate and similarly, there is a one - way causality from export to exchange rate.

4.3.4.4 Model Diagnostics

In this section, we perform the model diagnostic test discussed in section 2.3.5. which is crucial in the model acceptance phase before any good inference can be drawn. First the residuals of the variables show a realization of white noise process. (See Figure 4.7). We then consider theDurbin - Watson statistical test to detect the presence of

Table 4.6: Short run disequilibrium dynamics of Vector Error Correction Model

Independent Variables	$\Delta LEXP$ statistic	p value	$\Delta LGDP$ statistic	p value	$\Delta LEXC$ statistic	pvalue
Constant	0.009297	0.1918	0.060570**	0.0000	0.042055**	0.0000
$\Delta LEXP_{t-1}$	- 0.186315**	0.0000	0.097228**	0.0003	-0.126256**	0.0001
$\Delta LEXP_{t-2}$	- 0.67330**	0.00293	0.072762**	0.0072	- 0.058430	0.0652
$\Delta LGDP_{t-1}$	- 0.016202	0.6675	0.150224**	0.0000	- 0.096422**	0.0129
$\Delta LGDP_{t-2}$	- 0.013152	0.7233	0.046910	0.1582	0.01030	0.1897
$\Delta LEXC_{t-1}$	0.034731	0.27776	0.009032	0.7456	0.322404**	0.0000
$\Delta LEXC_{t-2}$	- 0.009628	0.7605	0.011206	0.6788	0.090138**	0.0049

Notes: ** means significant at 5% and 10% level of significance

Table 4.7: Pairwise Granger causality test

Wald Test - (Ch-square estimates)			
Dependent Variables	Inependent Variables	chi - squared value	p value
LEXP	$\Delta LGDP$	0.350740	0.839100
	$\Delta LEXC$	1.189554	0.551700
LGDP	ΔEXP	17.98342	0.0001**
	$\Delta LEXC$	0.205929	0.902200
LEXC	$\Delta LGDP$	7.082855	0.0290**
	$\Delta LEXP$	17.96487	0.0001**

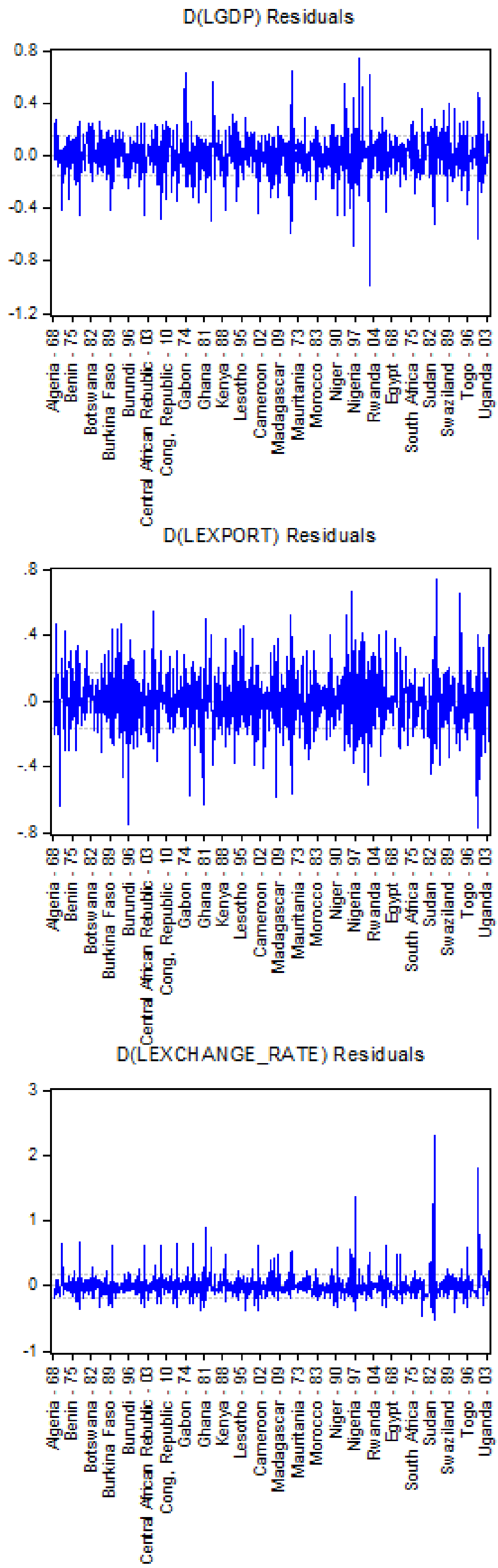
Notes: ** means significant at 5% and 10% level of significance. Restrictions are linear in coefficients.

autocorrelation in the residuals, table 4.8 shows the results of the test procedure. From table 4.8, we conclude that no serial correlation exist between the residuals of the first difference of our variables, since the test value in all the three cases is about $d = 2$. The residual plot can be found on appendix C .2

Table 4.8: Durbin - Watson Test

Dependent model Variables	Statistic value
$\Delta LEXP$	2.034610
$\Delta LGDP$	2.021981
$\Delta LEXC$	2.066931

Figure 4.7: Residual plot of the first difference



4.4 Country - by - country Analysis

In this section we follow the same procedure as we did in section 4.3. The only difference here is that we perform the granger causality for country by country. Instead of using the VAR model we decomposed our system model into individual equaltions and solve the system of equations using the Seemingly Unrelated Regression (SUR) model as discussed in section. Due to the fact that our trivariate vaariables are all cointegrated of order 1, we performed the Johansen - Fisher Panel Cointegration Test for the individual countries and the results can be found in appendix B .1.

4.4.1 Countrywise VEC models

Based on the Fisher - Johansen test results in appendix B, we estimate an SUR model which provides an effective way to analyze the disequilebrium dynamics our the variables both in the short run and in the long run. The results are illustrated in table 4.9. From the results represented in table 4.9, the long rum VEC model for each country can be represented by taking the negatives of each of the estimated coefficients given as:

$$LGDP_{t-1} = -\alpha_i - \beta_i LEXP_{t-1} - \gamma_i LEXC_{t-1} \tag{4.4}$$

where α_i is the constant term on table 4.9, β_i the estimated value of LGDP and γ_i the estimated value of LEXC from the table 4.9 respectively. For instance, the long run VEC model for Algeria will be given as:

$$LGDP_{t-1} = 0.267277 + 0.110173 LEXP_{t-1} + 0.148503 LEXC_{t-1} \tag{4.5}$$

Following the same analysis aas we did in section 4.3.3, we can see that there is a positve correlation between gdp and export in the following countries *(Botswana, Burkina, Burundi, Ghana, Lesotho , Niger, Egypt, South Africa, Togo and Uganda)*, whilst there is a positive correlation between exchange rate and export in the following countries *Cameroon, Rwanda and Sudan.* Moreover, we also notice that, there is a positive correlation between both gdp and exchange rate jointly on export in these countries, *(Algeria, Benin, Congo, Cote D'Ivoire, and Gabon.)* whilst there is a jointly negative correlation between gdp and exchange rate on export in these countries: *Central African Republic, Kenya and Malawi.*

NB:

Similar analysis can be performed for tables 4.10 and 4.11

Table 4.9: Vector Error Correction Estimates for each country.

Dependent variable: EXPORT

Country	Cointegration Eqn	Est. value	t-statistic
Algeria	LEXP	1.000000	
	LGDP	-0.110173	[-1.99833]
	LEXC	-0.148503	[-3.16451]
	CONSTANT	-0.267277	
Benin	LEXP	1	
	LGDP	-0.124095	[-1.81685]
	LEXC	-0.275299	[-1.73202]
	CONSTANT	1.353232	
Botswana	LEXP	1	
	LGDP	-0.286015	[-6.32939]
	LEXC	0.316038	[3.30338]
	CONSTANT	1.988146	
Burkina Faso	LEXP	1	
	LGDP	-0.38032	[-4.76311]
	LEXC	0.312352	[-4.76311]
	CONSTANT	4.095461	
Burundi	LEXP	1	
	LGDP	-0.135278	[-2.85726]
	LEXC	0.224941	[7.19213]
	CONSTANT	-0.665657	
Central African Rebublic	LEXP	1	
	LGDP	0.177465	[0.91922]
	LEXC	1.399815	[3.57242]
	CONSTANT	-14.88322	
Cong, Republic	LEXP	1	
	LGDP	-0.15961	[-4.12244]
	LEXC	-0.483859	[-4.78091]
	CONSTANT	2.275611	
Cote d'Ivoire	LEXP	1	
	LGDP	-0.198904	[-2.24178]
	LEXC	-0.221573	[-1.07997]
	CONSTANT	2.161382	
Gabon	LEXP	1	
	LGDP	-0.016001	[-0.40418]
	LEXC	-0.043304	[-0.38143]
	CONSTANT	-3.395155	
Ghana	LEXP	1	
	LGDP	0.358762	[2.02586]

	LEXC	-0.156734	[-5.29294]
	CONSTANT	-11.69068	
Kenya	LEXP	1	
	LGDP	0.068948	[1.87723]
	LEXC	0.028945	[0.89927]
	CONSTANT	-4.931922	
Lesotho	LEXP	1	
	LGDP	-0.801012	[-5.01202]
	LEXC	0.35109	[2.05724]
	CONSTANT	12.43526	
Cameroon	LEXP	1	
	LGDP	-0.013202	[-0.39870]
	LEXC	0.081314	[0.99705]
	CONSTANT	-3.28691	
Madagascar	LEXP	1	
	LGDP	0.077162	[1.01785]
	LEXC	-0.185366	[-7.13149]
	CONSTANT	-3.561213	
Malawi	LEXP	1	
	LGDP	0.008129	[0.11179]
	LEXC	0.046266	[1.43130]
	CONSTANT	-3.411333	
Mauritania	LEXP	1	
	LGDP	-0.007778	[-0.07337]
	LEXC	0.122348	[1.05526]
	CONSTANT	-4.064884	
Morocco	LEXP	1	
	LGDP	-0.303033	[-2.93629]
	LEXC	0.968915	[2.77569]
	CONSTANT	2.180012	
Niger	LEXP	1	
	LGDP	-0.186059	[-1.24770]
	LEXC	0.327247	[1.63327]
	CONSTANT	-0.789347	
Nigeria	LEXP	1	
	LGDP	-0.715866	[-1.12734]
	LEXC	0.634088	[2.45205]
	CONSTANT	13.01758	
Rwanda	LEXP	1	

	LGDP	-0.120226	[-1.23235]
	LEXC	0.021289	[0.16463]
	CONSTANT	0.172934	
Senegal	LEXP	1	
	LGDP	0.025519	[0.59756]
	LEXC	-0.06371	[-0.81881]
	CONSTANT	-3.461322	
Egypt	LEXP	1	
	LGDP	-0.467263	[-3.49652]
	LEXC	0.343671	[2.74142]
	CONSTANT	8.190869	
South Africa	LEXP	1	
	LGDP	-0.469527	[-3.34253]
	LEXC	0.469985	[3.31193]
	CONSTANT	8.143117	
Sudan	LEXP	1	
	LGDP	-0.195554	[-1.34385]
	LEXC	0.057998	[1.51769]
	CONSTANT	2.458167	
Swaziland	LEXP	1	
	LGDP	0.042716	[0.80061]
	LEXC	-0.062767	[-1.00688]
	CONSTANT	-5.059893	
Togo	LEXP	1	
	LGDP	-0.398843	[-2.12211]
	LEXC	1.147052	[3.26350]
	CONSTANT	-2.167984	
Uganda	LEXP	1	
	LGDP	-0.378199	[-1.70168]
	LEXC	0.09562	[2.69902]
	CONSTANT	5.498134	

Table 4.10: Vector Error Correction Estimates for each country (dependent variable GDP).

Dependent variable: GDP			
Country	Cointegration Eqn	Est. value	t-statistic

Algeria	LGDP	1	
	LEXP	-9.07661417	[-4.17852]
	LEXC	1.34790115	[2.75393]
	CONSTANT	2.425968277	
Benin	LGDP	1.000000	
	LEXP	-8.058314	[-5.15507]
	LEXC	2.218448	[1.92694]
	CONSTANT	-10.90477	
Botswana	LGDP	1.000000	
	LEXP	-3.496315	[-7.41851]
	LEXC	-1.104969	[-5.89017]
	CONSTANT	-6.951185	
Burkina Faso	LGDP	1	
	LEXP	-2.629366	[-4.07882]
	LEXC	-0.821287	[-1.99792]
	CONSTANT	-10.76847	
Burundi	LGDP	1	
	LEXP	-7.392169	[-4.81818]
	LEXC	-1.662799	[-4.71292]
	CONSTANT	4.920652	
Central African Rebublic	LGDP	1	
	LEXP	5.634929	[1.70508]
	LEXC	7.88786	[3.56322]
	CONSTANT	-83.86591	
Cong, Republic	LGDP	1	
	LEXP	-6.265273	[-5.41016]
	LEXC	3.031511	[3.03574]
	CONSTANT	-14.2573	
Cote d'Ivoire	LGDP	1	
	LEXP	-5.027552	[-1.44497]
	LEXC	1.113968	[0.84497]
	CONSTANT	-10.86646	
Gabon	LGDP	1	
	LEXP	-62.49429	[-3.50891]
	LEXC	2.706261	[0.43956]
	CONSTANT	212.1778	
Ghana	LGDP	1	
	LEXP	2.78736	[5.91602]
	LEXC	-0.436874	[-7.00012]

	CONSTANT	-32.58613	
Kenya	LGDP	1	
	LEXP	14.50362	[5.79654]
	LEXC	0.41981	[1.38187]
	CONSTANT	-71.53071	
Lesotho	LGDP	1	
	LEXP	-1.248421	[-3.19745]
	LEXC	-0.438309	[-2.07303]
	CONSTANT	-15.52445	
Cameroon	LGDP	1	
	LEXP	-75.74591	[-4.73537]
	LEXC	-6.15922009	[-1.21551]
	CONSTANT	248.9699677	
Madagascar	LGDP	1	
	LEXP	12.95976	[4.67151]
	LEXC	-2.402293	[-4.81877]
	CONSTANT	-46.15247	
Malawi	LGDP	1	
	LEXP	123.0197	[2.89070]
	LEXC	5.691616	[1.88536]
	CONSTANT	-419.6612	
Mauritania	LGDP	1	
	LEXP	-0.063572	[-0.12113]
	LEXC	8.173392	[4.19884]
	CONSTANT	-33.22389	
Morocco	LGDP	1	
	LEXP	-3.29997	[-1.99445]
	LEXC	-3.19739	[-3.04197]
	CONSTANT	-7.193974	
Niger	LGDP	1	
	LEXP	-5.374639	[-3.93620]
	LEXC	-1.758837	[-1.98652]
	CONSTANT	4.242454	
Nigeria	LGDP	1	
	LEXP	-1.396909	[-0.64553]
	LEXC	-0.885763	[-1.85850]
	CONSTANT	-18.18437	
Rwanda	LGDP	1	
	LEXP	-8.317652	[-3.22534]

	LEXC	-0.177074	[-0.19845]
	CONSTANT	-1.438408	
Senegal	LGDP	1	
	LEXP	39.18616255	[5.19314]
	LEXC	-2.4965508	[-1.04867]
	CONSTANT	-135.635937	
Egypt	LGDP	1	
	LEXP	-2.140123	[-3.99589]
	LEXC	-0.735499	[-6.23891]
	CONSTANT	-17.52947	
South Africa	LGDP	1	
	LEXP	-2.129803	[-2.07312]
	LEXC	-1.000975	[-7.92221]
	CONSTANT	-17.34324	
Sudan	LGDP	1	
	LEXP	-5.113684	[-4.33117]
	LEXC	-0.296584	[-2.33022]
	CONSTANT	-12.57029	
Swaziland	LGDP	1	
	LEXP	23.41065	[4.72721]
	LEXC	-1.469418	[-2.54951]
	CONSTANT	-118.4554	
Togo	LGDP	1	
	LEXP	-2.507252	[-1.61521]
	LEXC	-2.875948	[-3.70611]
	CONSTANT	5.435683	
Uganda	LGDP	1	
	LEXP	-2.644108	[-3.49824]
	LEXC	-0.252829	[-4.83967]
	CONSTANT	-14.53766	

Table 4.11: Vector Error Correction Estimates for each country (dependent variable Exchange Rate).

Dependent variable: EXCHANGE RATE			
Country	Cointegration Eqn	Est. value	t-statistic
Algeria	LEXC	1	

	LGDP	0.741894	2.09098
	LEXP	-6.733887101	0.042333
	CONSTANT	1.799811712	
Benin	LEXC	1.000000	
	LGDP	0.450766	[1.81362]
	LEXP	-3.63241	[-4.62534]
	CONSTANT	-4.915493	
Botswana	LEXC	1	
	LGDP	-0.905003	[-7.73955]
	LEXP	3.164175	[5.08745]
	CONSTANT	6.290843	
Burkina Faso	LEXC	1	
	LGDP	-1.217601	[-3.02619]
	LEXP	3.201519	[2.41859]
	CONSTANT	13.1117	
Burundi	LEXC	1	
	LGDP	-0.601395534	[-3.16431]
	LEXP	4.445617299	[8.14290]
	CONSTANT	-2.95925841	
Central African Rebublic	LEXC	1	
	LGDP	0.126777	[0.53153]
	LEXP	0.71438	[0.98849]
	CONSTANT	-10.63228	
Cong, Republic	LEXC	1	
	LGDP	0.329869	[2.75063]
	LEXP	-2.066716	[-5.68504]
	CONSTANT	-4.703042	
Cote d'Ivoire	LEXC	1	
	LGDP	0.897692	[2.40383]
	LEXP	-4.513192	[-1.98035]
	CONSTANT	-9.754731	
Gabon	LEXC	1	
	LGDP	0.369514	[0.48783]
	LEXP	-23.09249	[-3.67506]
	CONSTANT	78.40256	
Ghana	LEXC	1	
	LGDP	-2.288987	[-3.72266]
	LEXP	-6.380233	[-8.21986]
	CONSTANT	74.58924	

Kenya	LEXC	1	
	LGDP	2.382031	[2.42395]
	LEXP	34.54806	[4.87079]
	CONSTANT	-170.3884	
Lesotho	LEXC	1	
	LGDP	-2.281497	[-5.49735]
	LEXP	2.848269	[3.48036]
	CONSTANT	35.41898	
Cameroon	LEXC	1	
	LGDP	-0.162358218	[-0.49005]
	LEXP	12.29797098	[4.77428]
	CONSTANT	-40.42232033	
Madagascar	LEXC	1	
	LGDP	-0.416269	[-1.28140]
	LEXP	-5.394745	[-8.70369]
	CONSTANT	19.21184	
Malawi	LEXC	1	
	LGDP	0.175697	[0.15399]
	LEXP	21.6142	[3.02295]
	CONSTANT	-73.73322	
Mauritania	LEXC	1	
	LGDP	-0.063572	[-0.12113]
	LEXP	8.173392	[4.19884]
	CONSTANT	-33.22389	
Morocco	LEXC	1	
	LGDP	-0.312755	[-2.92896]
	LEXP	1.032083	[1.81532]
	CONSTANT	2.249952	
Niger	LEXC	1	
	LGDP	-0.568558	[-1.22875]
	LEXP	3.055792	[3.18713]
	CONSTANT	-2.41208	
Nigeria	LEXC	1	
	LGDP	-1.128971	[-0.96973
	LEXP	1.577069	[0.73262]
	CONSTANT	20.52962	
Rwanda	LEXC	1	
	LGDP	-5.647362	[-1.50891]
	LEXP	46.97279	[3.27608]

	CONSTANT	8.123213	
Senegal	LEXC	1	
	LGDP	-0.400552634	[-0.77004]
	LEXP	-15.69612063	[-5.22523]
	CONSTANT	54.32933188	
Egypt	LEXC	1	
	LGDP	-1.359621	[-7.71530]
	LEXP	2.909756	[3.87433]
	CONSTANT	23.83343	
South Africa	LEXC	1	
	LGDP	-0.999026	[-7.89982]
	LEXP	2.12772	[2.04834]
	CONSTANT	17.32635	
Sudan	LEXC	1	
	LGDP	-3.371722	[-2.02894]
	LEXP	17.24192	[4.25903]
	CONSTANT	42.38352	
Swaziland	LEXC	1	
	LGDP	-0.680542	[-2.09075]
	LEXP	-15.93192	[-4.87537]
	CONSTANT	80.6138	
Togo	LEXC	1	
	LGDP	-0.347711	[-2.62862]
	LEXP	0.8718	[1.76179]
	CONSTANT	-1.890049	
Uganda	LEXC	1	
	LGDP	-3.955245	[-3.17638]
	LEXP	10.4581	[3.64161]
	CONSTANT	57.50001	

4.4.2 Long run country - by - country causality analysis:

In this section, we try to discover the joint causality between our three variables. From table 4.12, it can be seen that, GDP and Exchange rate jointly Granger cause export at 5% and 10% level of significance in the following countries: Algreria, Benin, Burkina Faso, Burundi, Central African Republic, Cabon, Kenya, Cameroon, Madagascar, Malawi, Mauritania, Morocco, niger, Sengal, Egypt, South africa, Swazilan and Uganda.

Table 4.12: Long run causality analysis

Granger Causality tests, trivate models H0 : LGDP and LEXC jointly does not cause LEXP			
Country	**Est. value**	**t - statistic**	**p value**
Algeria	-0.352399	-2.342803	0.01075*
Benin	-0.387829	-3.52176	0.0006*
Botswana	-0.221529	-1.559644	0.1219
Burkina Faso	-0.46113	-3.199131	0.0018*
Burundi	-0.8344	-3.068731	0.0027*
Central African Rebublic	-0.136807	-3.709406	0.0003*
Cong, Republic	-0.311857	-1.320415	0.1896
Cote d'Ivoire	-0.086999	-0.906049	0.367
Gabon	-0.504623	-2.954741	0.0039*
Ghana	-0.208035	-1.540705	0.1264
Kenya	-0.491488	-2.641172	0.0095*
Lesotho	-0.068771	-0.808587	0.4206
Cameroon	-0.804621	-5.365012	0.0001*
Madagascar	-0.62877	-3.102221	0.0025*
Malawi	-0.698174	-2.954467	0.0039*
Mauritania	-0.478046	-4.078032	0.0001*
Morocco	-0.167908	-2.696079	0.0082*
Niger	-0.27086	-3.670748	0.00038*
Nigeria	-0.045814	-1.690618	0.0939
Rwanda	-0.168359	-1.446647	0.151
Senegal	-0.554879	-4.15301	0.000066*
Egypt	-0.219782	-2.638395	0.0096*
South Africa	-0.149148	-2.167732	0.0324*
Sudan	0.004765	0.053759	0.9572
Swaziland	-0.500365	-3.716561	0.0003*
Togo	0.032875	0.393629	0.6947
Uganda	-0.319809	-3.136101	0.0022*

Notes: ** means significant at 5% and 10% level of significance. Restrictions are linear in coefficients.

From table 4.13, it can be observed that, Export and Exchange rate jointly Granger cause GDP in these countries: Algeria, Benin, Cote d'Ivoire, Kenya, Lesotho, madagascar, Malawi, Rwanda, Egypt, South Africa, Sudan, Swaziland, Togo and Uganda.

Table 4.13: Long run causality analysis

Granger Causality tests, trivate models H0 : LXP and LEXC jointly does not cause LGDP			
Country	**Est. value**	**t - statistic**	**p value**
Algeria	-0.024601997	-2.352037822	0.0205*
Benin	-0.387829	-3.52176	0.0056*
Botswana	-0.050656	-1.219777	0.2253
Burkina Faso	-0.055332	-1.172187	0.2438
Burundi	-0.033005	-1.843397	0.0681
Central African Rebublic	0.011801	1.591725	0.1145
Cong, Republic	-0.06631948	-1.591929447	0.11440677
Cote d'Ivoire	-0.042696	-2.147338	0.0341*
Gabon	-0.005313	-1.461012	0.147
Ghana	0.058032	1.745105	0.0839
Kenya	0.040844	3.239731	0.0016*
Lesotho	-0.122621	-2.493621	0.0142*
Cameroon	-0.002943309	-1.839329936	0.068692156
Madagascar	0.04047	3.190321	0.0019*
Malawi	0.004919743	2.33501974	0.021444*
Mauritania	0.006271	0.822328	0.4128
Morocco	-0.009842676	-0.456674824	0.648848101
Niger	-0.013419	-0.867393	0.3877
Nigeria	-0.039756	0.020649	0.0569
Rwanda	-0.036753	-2.989683	0.0035*
Senegal	0.006231306	1.603784135	0.111765532
Egypt	-0.095162557	-4.461027422	0.0000205*
South Africa	-0.088229095	-2.055159322	0.04234513*
Sudan	-0.054331	-5.01117	0.000001*
Swaziland	0.017724	1.994643	0.0487*
Togo	-0.070588	-2.774848	0.0065*
Uganda	-0.098615	-3.021352	0.0032*

Notes: ** means significant at 5% and 10% level of significance. Restrictions are linear in coefficients.

Finally, from table 4.14, GDP anad export jointly cause exchange rate in these countries: Algeria, Central Africa Republic, Ghana, Lesotho, Cameroon, Madagascar, Morocco, South Africa, Sudan and Swaziland.

Table 4.14: Long run causality analysis

Granger Causality tests, trivate models H0 : LGDP and LEXP jointly does not cause LEXC			
Country	**Est. value**	**t - statistic**	**p value**
Algeria	0.042330802	2.985093759	0.0035*
Benin	-0.03405	-1.145672	0.2545
Botswana	-0.017403	-0.417334	0.6773
Burkina Faso	-0.059228	-1.357981	0.1774
Burundi	-0.01786611	-0.80392753	0.423255
Central African Rebublic	-0.176125	-2.707754	0.0079*
Cong, Republic	-0.071856	-0.63655	0.5258
Cote d'Ivoire	0.032033	1.092366	0.2772
Gabon	-0.001322	-0.164967	0.8693
Ghana	0.095729	5.033938	0.00001*
Kenya	-0.006604	-1.132654	0.2599
Lesotho	-0.074944	-4.669871	0.00001*
Cameroon	-0.03070721	-2.13824612	0.034816*
Madagascar	0.133485	4.270096	0.00001*
Malawi	-0.01167	-1.213193	0.2278
Mauritania	0.006271	0.822328	0.4128
Morocco	-0.12146	-2.574625	0.0114*
Niger	-0.001566	-0.059074	0.953
Nigeria	0.008273	0.439554	0.6612
Rwanda	-0.002033	-1.886524	0.062
Senegal	0.011166644	0.996532796	0.3212823
Egypt	-0.084674	-2.854082	0.0052*
South Africa	-0.13959368	-3.49436668	0.00069*
Sudan	-0.021941	-2.7739	0.0066*
Swaziland	0.022396	2.03366	0.0445*
Togo	-0.15495	-0.823882	0.4119
Uganda	0.001553	0.101725	0.9192

Notes: ** means significant at 5% and 10% level of significance. Restrictions are linear in coefficients.

4.4.3 Pairwise Causlity, country - by - country

We now now ready to investigate the causality between our variables taking pairwise instead of jointly as we investigated in section 4.4.2. In this section all the test was performed using the Wald test and the results are illustrated in the following sections.

4.4.3.1 Causality from LGDP to LEXP and causality from LEXC to LEXP

From table 4.15, the Granger causality test for the null hypotheses LGDP does not cause LEXP and LEXC does not cause LEXP are shown in table 4.15. From table 4.15, we noticed that there was only 6 out of the 27 countries that provided significant evidence to reject the null hypothesis that LGDP does not cause LEXP, that is, these countries Central Africa Republic, Cameroon, Morocco, Senegal and Swaziland,provided sufficient evidence to conclude that GDP Granger cause export.

Also, we can see that, LEXC granger cause LEXP in these countries at 5% and 10% levels of significance. These countries are Benin, Cameroon, Morocco, Niger, Swaziland and Uganda.

Table 4.15: Short run disequilibrium dynamics of Vector Error Correction Model

	Granger Causality tests H0 : LGDP does not cause LEXP		Granger causality test H0: LEXC does not cause LEXP	
Country	Chi - squd. value	p value	Chi - squd. value	p value
Algeria	0.086971073	0.95744	0.939043	0.6253
Benin	4.863476	0.0879	6.287612	0.0431*
Botswana	0.009702	0.9952	0.129583	0.9373
Burkina Faso	2.88421	0.2364	3.069095	0.2156
Burundi	2.072322	0.3548	0.97064	0.6155
Central African Rebublic	10.05327	0.0066*	3.381155	0.1844
Cong, Republic	0.526013187	0.76873	0.152399569	0.9266311
Cote d'Ivoire	3.118464	0.2103	3.970688	0.1373
Gabon	1.365898	0.5051	0.119408	0.942
Ghana	0.817829	0.6644	0.086279	0.9578
Kenya	1.121361	0.5708	1.233009	0.5398
Lesotho	0.019423	0.9903	0.204888	0.9026
Cameroon	13.03783	0.0015*	27.92469	0.00001*
Madagascar	0.625516	0.7314	1.339033	0.512
Malawi	1.912529	0.3843	2.771635	0.2501
Mauritania	5.96243	0.0507	0.928995	0.6285
Morocco	7.918513	0.0191*	7.375953	0.025*

Niger	5.758232	0.0562	6.703767	0.035*
Nigeria	4.020813	0.1339	5.620732	0.0602
Rwanda	3.703573	0.157	0.105524	0.9486
Senegal	7.177717755	0.02762*	3.198873952	0.20201
Egypt	2.679279	0.2619	1.52826	0.4657
South Africa	7.037998	0.0296*	5.03539	0.0806
Sudan	0.531229	0.7667	0.250186	0.8824
Swaziland	6.836756	0.0328*	6.314697	0.0425*
Togo	0.037265	0.9815	0.382371	0.826
Uganda	0.85582	0.6519	12.09538	0.0024*

Notes: ** means significant at 5% and 10% level of significance. Restrictions are linear in coefficients.

4.4.3.2 Causality from LEXP to LGDP and causality from LEXC to LGDP

From table 4.16, the Granger causality test for the null hypotheses LEXP does not cause LGDD and LEXC does not cause LEXP are shown in table 4.16. From table 4.16, we noticed that there was only 3 out of the 27 countries that provided significant evidence to reject the null hypothesis that LEXP does not cause LGDP at 5% and 10% level of significance. These countries are: Madagascar, Sudan, and Uganda.

As we can see, Benin, Central Africa Republic, Cameroon, Mauritania, Egypt and Suadan provided sufficient evidence at 10% and 5% levels of significance that, LEXC causes LGDP

Table 4.16: Short run disequilibrium dynamics of Vector Error Correction Model

	Granger Causality tests H0 : LEXP does not cause LGDP		Granger causality test H0: LEXC does not cause LGDP	
Country	Chi - squad value	p value	Chi - squad value	p value
Algeria	0.081859186	0.9598	1.104556433	0.57563
Benin	0.530692	0.7669	10.51533	0.0052*
Botswana	0.372763	0.83	0.276796	0.8708
Burkina Faso	2.983235	0.225	2.983235	0.225
Burundi	1.339455	0.5118	0.955351	0.6236
Central African Rebublic	2.96494	0.2271	8.936915	0.0115*
Cong, Republic	0.834112383	0.6589	1.968449854	0.373728
Cote d'Ivoire	0.635335	0.7278	1.371988334	0.5035
Gabon	0.773361	0.6793	2.075828	0.3542
Ghana	0.0839	0.0832	4.654026	0.0976

Kenya	0.68832	0.7088	2.364758	0.3065
Lesotho	1.988996	0.3699	5.831433	0.0542
Cameroon	0.200768329	0.9045	18.86554593	0.00008*
Madagascar	6.169745	0.0457*	0.209033	0.9008
Malawi	1.410982	0.4939	2.974029	0.226
Mauritania	4.593519	0.1006	7.469698	0.0239*
Morocco	0.947455	0.6227	3.359349	0.1864
Niger	1.463316	0.4811	1.332018	0.5138
Nigeria	0.706022	0.7026	3.780125	0.1511
Rwanda	3.122643	0.2099	0.952916	0.6210
Senegal	0.144899757	0.93	3.782598878	0.1507
Egypt	0.215977	0.8976	10.20503	0.0061*
South Africa	0.345792	0.8412	1.73392	0.4202
Sudan	7.637562	0.022*	13.11227	0.0014*
Swaziland	2.454716	0.2931	1.748485	0.4172
Togo	0.92757	0.6289	2.717212	0.257
Uganda	9.510406	0.0086*	4.520435	0.1043

Notes: ** means significant at 5% and 10% level of significance. Restrictions are linear in coefficients.

4.4.3.3 Causality from LGDP to LEXC and causality from LEXP to LEXC

Finally, from table 4.17, we can observe that Benin, Burkina Faso, Central Africa Republic, Lesotho, Madagascr, Niger, Rwanda, Togo and Uganda, provide sufficient evidence to reject the null hypothesis, LGDP does not cause LEXC and accept the alternative hypothesis that, LGDP Granger cause LEXC.

Finally, out of all the 27 countries considered for this analysis, only 2 countries, Madagascar and Mauritania exihibit significant evidence to reject the null hypothesis that, LEXP does not cause LEXP and conclude that LEXP causes LEXC in thses countries.

Table 4.17: Short run disequilibrium dynamics of Vector Error Correction Model

	Granger Causality tests H0 : LGDP does not cause LEXC		Granger causality test H0: LEXP does not cause LEXC	
Country	**Chi - squared value**	**p value**	**Chi - squared value**	**p value**
Algeria	1.276237285	0.52828	1.198046689	0.54934
Benin	16.02105	0.0003*	2.414176	0.2991
Botswana	0.593516	0.7432	1.656324	0.4369

Burkina Faso	8.293527	0.0158*	0.07380562	0.9637698
Burundi	0.937700172	0.62572138	0.168577841	0.9191
Central African Rep.	7.908678	0.0192*	1.619792	0.4449
Cong, Republic	2.971714	0.2263	0.20046	0.9046
Cote d'Ivoire	0.080459	0.9606	2.392272	0.3024
Gabon	1.016461	0.6016	0.168743	0.9191
Ghana	3.018201	0.2211	5.289442	0.071
Kenya	0.093771	0.9542	0.177554	0.915
Lesotho	11.93732	0.0026*	0.496749	0.7801
Cameroon	4.426519884	0.1093436	0.812404901	0.66617
Madagascar	6.821003	0.033*	12.11018	0.0023*
Malawi	0.273831	0.872	1.693178	0.4289
Mauritania	4.593519	0.1006	7.469698	0.0239*
Morocco	0.334737	0.8459	3.093188	0.213
Niger	21.05828222	0.000026*	1.376425182	0.5024
Nigeria	3.207064	0.2012	4.356231	0.1133
Rwanda	19.85313	0.00001*	1.208906	0.5464
Senegal	4.33450277	0.1144	0.543756973	0.7619
Egypt	0.73389	0.6928	5.774107	0.0557
South Africa	0.108358	0.9473	1.060944	0.5883
Sudan	4.461815	0.1074	0.062707	0.9691
Swaziland	0.109774	0.9466	2.619025	0.27
Togo	13.64457	0.0011*	0.875623	0.6454
Uganda	13.29358	0.0013*	2.567362	0.277

Notes: ** means significant at 5% and 10% level of significance. Restrictions are linear in coefficients.

Chapter 5

CONCLUSION

5.1 Summary

This thesis investigates the possibility of Granger causality between the natural logarithms of GDP, Export and Exchange rate in 27 developing countries in African from 1965 to 2010. A panel data approach based on SUR systems and the WALD test with country specify testing was applied to the variables. Two differents models have been studied. A trivariate (GDP - Export - Exchange rate) models, one for the overall panel data and the second based on country - by - country model. All things considerd, all the three variables passed the unit root test and Johansen procedures indicated a statistical significance between the variables.

We noticed that, in the long run, there is a positive correlation between GDP and Eport and a negative correlation between export and exchange rate. Finally, there is a positive between relationship between GDP and exchange rate in the long run. Also, we observed that, there was no joint causality from GDP and exchange rate to Export. Similarly, there was no joint causality causality from export and exchange rate to GDP, however, there is a unidirectional causality from GDP and export to exchange rate.

For the pairwise causality analysis,we noticed that there was no evidence of causality from either GDP or exchange rate to export but there was a unidirectional causality fronn GDP to exchange rate and similarly, there was a one - way causality from export to exchange rate.

For the country - by -country analysis, our results indicate a jointly causality from GDP and exchange rate to export in Algreria, Benin, Burkina Faso, Burundi, Central African Republic, Cabon, Kenya, Cameroon, Madagascar, Malawi, Mauritania, Morocco, niger, Sengal, Egypt, South africa, Swazilan and Uganda. Similarly, there is ajoint causality from export and exchange rate to GDP in Algeria, Benin, Cote d'Ivoire, Kenya, Lesotho, Madagascar, Malawi, Rwanda, Egypt, South Africa, Sudan, Swaziland, Togo and Uganda. Finally, GDP and export jointly Granger cause exchange rate in

Algeria, Central Africa Republic, Ghana, Lesotho, Cameroon, Madagascar, Morocco, South Africa, Sudan and Swaziland.

We also noticed that there is a one - causality from GDP to Export in Central Africa Republic, Cameroon, Morocco, Senegal and Swaziland, one - way causality from export to GDP in Madagascar, Sudan, and Uganda , undirectional causality from exchange rate to export in Benin, Cameroon, Morocco, Niger, Swaziland and Uganda ,unidirecyional causality from export to exchange rate in Madagascar and Mauritania. Also there was a one - way causality from DGP to exchange rate in Burkina Faso, , Lesotho, Madagascar, Niger, Rwanda, Togo and Uganda and one way - one causality from exchange rate to export in Cameroon, Mauritania, Egypt and Sudan. We noticed that, there was a bidirectional or 2 - way causality from GDP to exchange rate in Benin, Central Africa Republic, whilst in case of Lgeria, Botswana, Burundi, Congo Republic, Cote d'Ivoire, Gabon, Ghana, Kenya, Lesotho, Malawi, Nigeria and South Africa there is no evidence of causality beteen these variables.

REMARK:

Finally, it is very important to state that Granger causality between exports and GDP does not necessarily mean that the Export - led Growth (ELG) or Growth - driven export (GDE) hypothesis is valid. The signs of the regression coefficients involved in the causality tests are alsovery crucial since the ELG and GDE hypotheses imply positive effects, that is, the parameters of our models are all expected to be *(*positive).

Bibliography

[1] Bahamni- Oskooee, M., H. Mohtadi, and G. Shabsign (1991), *Exports, Growth and Causality in LDCs: A Reexamination*, Journal of Development Economics, 36, 405-415.

[2] Bahamni- Oskooee, M., H. Mohtadi, and J. Alse (1993), *Export Growth and Economic Growth: An Application of Cointegration and Error-Correction Modeling*, Journal of Developing Areas, 27, 535-542.

[3] Bhagwati, J. (1979), *International Factor Movements and National Advantage*, 9th V.K. Ramawami Memorial Lecture

[4] Buffie E. (1992), *On the condition for export - led growth.* Canadian Journal of Economics, 25; 211 - 225.

[5] Balassa B.(1978), *Exports and economic growth: further evidence.* Journal of Development Economics, 5; 181-189.

[6] Bhagwati J.N. (1988) *Protectionism.* MIT Press, Cambridge. Massachusetts.

[7] Breusch, T., Pagan, A., 1980. *The LM Test and Its Applications to Model Specification in Econometrics*, Review of Economic Studies 47, 239-254.

[8] Breuer, J.B., McNown, R., Wallace, M., 2001. *Misleading Inferences from Panel Unit Root Tests with an Illustration from Purchasing Power Parity*, Review of International Economics 9, 482-493.

[9] Chow, P.C.Y. (1987), *Causality between Export Growth and Industrial Development*, Journal of Development Economics, 26, 55-63.

[10] Chien - Hui Lee and Bwo - Nung Haung (2002), *The Relationship Between Exports and Economic Growth in East Asian Countries: A Multivariate Threshold Autoregressive Approach*, Journal of Economics, vol. 27

[11] Dufour, J.M., Renault, E., 1998. *Short Run and Long Run Causality in Time Series:* Theory, Econometrica 66, 1099-1125.

[12] Diebold, F. X. & Kilian, L. (2000), *Unit- Root Tests are Useful for Selecting Forecasting Models*, Journal of Business & Economic Statistics 18(3), 265–273.

[13] Davidson, Russell; MacKinnon, James G. (2004). *Econometric Theory and Methods.* New York: Oxford University Press. p. 623.

[14] Davidson, James (2000). *Econometric Methods.* Wiley

[15] Edwards S.(1998), *Openness, productivity and growth: what do we really know?*,Economic Journal, 108; 383 - 398.

[16] Enders, W., 2004. *Applied Econometric Time Series*, 2nd ed.Wiley.

[17] Engle, R. F. & Granger, C. W. J. (1987), *Co-Integration and Error Correction: Representation, Estimation, and Testing*, Econometrica: journal of the Econometric Society 55(2), 251–276.

[18] Feder G.(1982), *On exports and economic growth.* Journal of Development Economics, 12; 59 - 73.

[19] Findlay R.(1984), *Growth and Development in Trade Models.* in Handbook of International Economics, vol. 1, (eds). Jones, R. and Kenen, P.. Amsterdam: North - Holland.

[20] Ghartey, E.E. (1993), *Causal Relationship between Exports and Economic Growth: Some Empirical Evidence in Taiwan, Japan and the U.S.*, Applied Economics, 25, 1145-1152.

[21] Grossman G, Helpman E. (1991) *Innovation and Growth in the Global Economy.* Cambridge, MA: MIT Press.

[22] Giles, J.A., Williams, C.L., 2000. *Export-Led Growth: A Survey of the Empirical Literature and Some Noncausality Results*, Part 1, Journal of International Trade and Economic Development 9, 261-337.

[23] Greene, W.H., 2003. *Econometric Analysis*, 2nd ed. Prentice-Hall.

[24] Granger, C.W.J., 1969. *Investigating causal relations by econometric models and cross - spectral methods.* Econometrica, 37, 424-438.

[25] Granger, C. W. (1986), *Developments in the Study of Cointegrated Economic Variables*, Oxford Bulletin of economics and statistics 48(3), 213–228.

[26] Huang, B.N., J.J.R. Oh, and C.W. Yang (2000), *On the Causality of Exports and Economic Growth: An Application of Direct and Indirect Causality*, Working Paper.

[27] Helpman E, Krugman P.(1985), *Market Structure and Foreign Trade.* Cambridge, MA: MIT Press.

[28] Jung, W., and P. Marshall (1985), *Exports, Growth and Causality in Developing Countries*, Journal of Development Economics 18, 1-12.

[29] Johansen, Søren (1991). *Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models.* Econometrica. 59 (6): 1551–1580

[30] Krugman P.R. (1984), *Import protection as export promotion. In: Kierzkowski, H. (Ed.), Monopolistic Competition in International Trade.* Oxford University Press, Oxford.

[31] Konya, L, 2006. *Exports and growth: Granger causality analysis on OECD countries with a panel data approach.* Econ. Model. 23(6), 978-992.

[32] Kirchgassner, G., Wolters, J. & Hassler, U. (2012),*Introduction to Modern Time Series Analysis*, Springer Science & Business Media.

[33] Lutkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer Science & Business Media.

[34] Michaely, M. (1977), *Exports and Growth: An Empirical Investigation*, Journal of Development Economics, 4(1), March, 49-53.

[35] Michalopoulos, C., and K. Jay (1973), *Growth of Exports and Income in the Developing World: A Neoclassical View*, A.I.D. Discussion Paper No. 28 (Washington, DC).

[36] Marin D. (1992) *Is the export - led growth hypothesis valid for industrialized countries?*, Review of Economics and Statistics, 74; 678 - 688.

[37] Oxley L. (1993), *Cointegration, causality and export - led growth in Portugal, 1865 - 1985.* Economic Letters, 43; 163 - 166.

[38] Pesaran, M.H., Yamagata, T., 2008. *Testing slope homogeneity in large panels.* J. Econometrics. 142, 50–93.

[39] Phillips, P. C. B.; Perron, P. (1988). *Testing for a Unit Root in Time Series Regression.* Biometrika. 75 (2): 335 - 346

[40] Riezman, R.G., C.H. Whiteman, and P.M. Summers (1996), *The Engine of Growth or Its Handmaiden? A Time Series Assessment of Export-Led Growth*, Empirical Economics, Vol. 21, No. 1, 77-110

[41] Rivera - Batiz L, Romer P.(1991) *Economic integration and endogenous growth.* Journal of Economics, 106; 531 - 556.

[42] Romer P.(1990) *Endogenous technological change.* Journal of Politic Economics, 98; 71 - 102.

[43] Ribeiro Ramos F.F,(2001), *Exports, imports, and economic growth in Portugal: evidence from causality and cointegration analysis.* Economic Modelling, 18; 613 - 623.

[44] Sharma, S.C., and D. Dhakal (1994), *Causal Analyses Between Exports and Economic Growth in Developing Countries*, Applied Economics, 26, 1145-1157.

[45] Sharma, S.C., M. Norris, and D.W. Cheung (1991), *Exports and Economic Growth in Industrialized Countries*, Applied Economics, 23, 697-708

[46] Shan, J., and F. Sun (1998), *On the Export - Led Growth Hypothesis for the Little Dragons: An Empirical Reinvestigation*, Atlantic Economic Journal, 26(4), Dec ., 353-71.

[47] Segerstrom P, Anant T, Dinopoulos E. (1990), *A Shumpeterian model of the product life cycle.* American Economic Review, 80; 1077 - 1091.

[48] Sims, C. A. (1980), *Macroeconomics and Reality*, Econometrica: Journal of the Econometric Society 48(1).

[49] Said, S. E. & Dickey, D. A. (1984), *Testing for Unit Roots in Autoregressive - Moving Average Models of Unknown Order*, Biometrika 71(3), 599 - 607

[50] Thornton J. (1996) *Cointegration, causality and export - led growth in Mexico.* Economic Letters, 50; 413 - 416

[51] Vernon R.(1996), *International investment and international trade in the product cycle.* Journal of Econometric Society Monograph, Cambridge, Cambridge University Press, 80; 190 - 207.

[52] Xu, Z. (1996), *On the Causality between Export Growth and GDP Growth: An Empirical Reinvestigation*, Review of International Economics, 4(2), 172-84.

[53] Zellner, A., 1962. *An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias*, Journal of the American Statistical Association 57, 348- 368

Appendix A: Summary Statistics

Table 1: Summary statistics for the various variables

	LEXCHANGE_RATE	LEXPORT	LGDP
Mean	3.236669419	3.146690603	21.92006255
Median	4.471638794	3.163473821	21.75648818
Maximum	7.685959118	4.614615499	26.65112328
Minimum	-9.547330389	1.204480424	17.63954725
Std. Dev.	3.405301254	0.625028369	1.690824787
Skewness	-1.406542136	-0.265590965	0.262238126
Kurtosis	5.203508742	2.809303338	2.842782394
Jarque-Bera	660.2582217	16.47011044	15.50178162
Probability	0	0.000265192	0.000430359
Sum	4016.706749	3905.043038	27202.79762
Sum Sq. Dev.	14379.13502	484.4189729	3545.021689
Observations	1241	1241	1241

.1 Appendix B: Johansen Fisher Panel Cointegration Test for each country

Johansen Fisher Panel Cointegration Test

Series: LGDP LEXCHANGE_RATE LEXPORT

Trend assumption: Linear deterministic trend (restricted)

Lags interval (in first differences): 1 2

Unrestricted Cointegration Rank Test (Trace and Maximum Eigenvalue)

Hypothesized No. of CE(s)	Fisher Stat.* (from trace test)	Prob.	Fisher Stat.* (from max-eigen test)	Prob.
None	122.1	0.0000	102.7	0.0001
At most 1	56.11	0.3956	45.64	0.7837
At most 2	43.28	0.8516	43.28	0.8516

*** Probabilities are computed using asymptotic Chi-square distribution.**

Individual cross section results

Cross Section	Trace Test Statistics	Prob.**	Max-Eign Test Statistics	Prob.**
Hypothesis of no cointegration				
Algeria	33.5197	0.3109	19.0227	0.3038
Benin	44.7631	0.0323	21.0757	0.1873
Botswana	33.7665	0.2993	15.5589	0.5842

Lags interval (in first differences): 1 1

Unrestricted Cointegration Rank Test (Trace and Maximum Eigenvalue)

Hypothesized No. of CE(s)	Fisher Stat.* (from trace test)	Prob.	Fisher Stat.* (from max-eigen test)	Prob.
None	100.1	0.0001	93.09	0.0008
At most 1	43.64	0.8421	41.75	0.8880
At most 2	30.04	0.9966	30.04	0.9966

*** Probabilities are computed using asymptotic Chi-square distribution.**

Individual cross section results

Cross Section	Trace Test Statistics	Prob.**	Max-Eign Test Statistics	Prob.**
Hypothesis of no cointegration				
Algeria	34.7080	0.2573	23.0982	0.1100
Benin	34.5171	0.2655	20.9167	0.1948
Botswana	30.3455	0.4819	18.2077	0.3615

Burkina Faso	24.7885	0.8003	11.8737	0.8804	Burkina Faso	29.3085	0.5434	17.2572	0.4363
Burundi	37.1358	0.1678	20.8020	0.2004	Burundi	33.7692	0.2992	19.9306	0.2471
Central African Rebublic	35.5377	0.2236	18.5510	0.3364	Central African Rebublic	33.3966	0.3169	16.8055	0.4742
Cong, Republic	27.2803	0.6648	12.0897	0.8671	Cong, Republic	34.9815	0.2459	20.9255	0.1944
Cote d'Ivoire	25.6032	0.7587	12.2546	0.8566	Cote d'Ivoire	25.1322	0.7832	9.7392	0.9699
Gabon	33.6488	0.3048	14.9964	0.6348	Gabon	33.9964	0.2886	14.3092	0.6958
Ghana	41.5321	0.0683	29.0394	0.0182	Ghana	42.3249	0.0572	29.4722	0.0158
Kenya	51.3234	0.0059	36.5926	0.0013	Kenya	43.3991	0.0447	25.6812	0.0522
Lesotho	55.6734	0.0017	33.2613	0.0043	Lesotho	34.1715	0.2807	18.3327	0.3522
Cameroon	42.8741	0.0505	27.6900	0.0281	Cameroon	41.0281	0.0763	29.2368	0.0171
Madagascar	38.9037	0.1190	21.2742	0.1781	Madagascar	48.8676	0.0114	29.9106	0.0136
Malawi	28.3586	0.6006	19.1909	0.2926	Malawi	42.1825	0.0591	23.7037	0.0929
Mauritania	25.6562	0.7559	14.5608	0.6736	Mauritania	25.3087	0.7741	13.6659	0.7506
Morocco	49.1318	0.0106	30.6224	0.0107	Morocco	51.7739	0.0052	35.1093	0.0023
Niger	32.9734	0.3377	14.3887	0.6888	Niger	31.1928	0.4330	17.5340	0.4138
Nigeria	37.8453	0.1467	19.6105	0.2662	Nigeria	31.0233	0.4427	17.3130	0.4317
Rwanda	36.5266	0.1877	20.5105	0.2151	Rwanda	45.2809	0.0284	23.8584	0.0889
Senegal	44.7144	0.0327	25.7981	0.0504	Senegal	35.7944	0.2139	16.6610	0.4866
Egypt	41.5773	0.0676	21.0193	0.1899	Egypt	38.6724	0.1247	19.3119	0.2848
South Africa	33.3837	0.3175	14.0001	0.7225	South Africa	37.5293	0.1558	18.4034	0.3471
Sudan	59.2963	0.0006	28.6234	0.0208	Sudan	54.0152	0.0028	23.8629	0.0888
Swaziland	36.6440	0.1838	20.3843	0.2218	Swaziland	25.7295	0.7520	15.7873	0.5637
Togo	22.9020	0.8816	11.0238	0.9252	Togo	26.9078	0.6864	14.5062	0.6784
Uganda	46.9778	0.0186	26.3863	0.0421	Uganda	31.6307	0.4086	17.3573	0.4281
Hypothesis of at most 1 cointegration relationship					**Hypothesis of at most 1 cointegration relationship**				
Algeria	14.4970	0.6164	7.9081	0.8296	Algeria	11.6098	0.8381	6.1615	0.9487
Benin	23.6873	0.0913	16.9059	0.1106	Benin	13.6004	0.6904	9.5617	0.6656
Botswana	18.2077	0.3301	10.5255	0.5629	Botswana	12.1378	0.8024	6.3364	0.9402
Burkina Faso	12.9148	0.7448	8.5885	0.7660	Burkina Faso	12.0514	0.8085	7.7093	0.8467
Burundi	16.3338	0.4660	12.6659	0.3557	Burundi	13.8386	0.6709	11.8834	0.4261

Central African Rebublic	16.9867	0.4160	14.1836	0.2421	Central African Rebublic	16.5911	0.4460	13.3489	0.3007
Cong, Republic	15.1906	0.5587	11.4501	0.4680	Cong, Republic	14.0559	0.6530	10.0047	0.6184
Cote d'Ivoire	13.3486	0.7107	8.0646	0.8157	Cote d'Ivoire	15.3930	0.5420	9.1278	0.7112
Gabon	18.6524	0.3018	12.9469	0.3324	Gabon	19.6872	0.2422	13.6433	0.2789
Ghana	12.4927	0.7767	9.5403	0.6679	Ghana	12.8526	0.7496	9.8556	0.6343
Kenya	14.7308	0.5970	10.6017	0.5548	Kenya	17.7180	0.3632	13.5702	0.2842
Lesotho	22.4120	0.1271	14.6738	0.2120	Lesotho	15.8388	0.5055	10.7293	0.5414
Cameroon	15.1841	0.5592	8.1478	0.8081	Cameroon	11.7913	0.8262	6.4030	0.9367
Madagascar	17.6295	0.3694	11.7239	0.4413	Madagascar	18.9571	0.2833	12.4690	0.3728
Malawi	9.1678	0.9555	4.7807	0.9897	Malawi	18.4788	0.3126	13.4739	0.2913
Mauritania	11.0954	0.8697	7.8783	0.8322	Mauritania	11.6428	0.8360	10.7556	0.5387
Morocco	18.5094	0.3107	12.6446	0.3575	Morocco	16.6645	0.4404	12.2008	0.3967
Niger	18.5848	0.3060	11.3395	0.4790	Niger	13.6588	0.6856	8.4666	0.7779
Nigeria	18.2348	0.3284	9.7576	0.6447	Nigeria	13.7103	0.6814	8.6724	0.7576
Rwanda	16.0160	0.4913	8.4131	0.7831	Rwanda	21.4225	0.1622	14.9709	0.1952
Senegal	18.9163	0.2857	10.8868	0.5250	Senegal	19.1333	0.2730	12.8926	0.3368
Egypt	20.5580	0.1989	16.3893	0.1294	Egypt	19.3605	0.2600	13.2386	0.3092
South Africa	19.3837	0.2587	13.5139	0.2884	South Africa	19.1260	0.2734	15.7510	0.1562
Sudan	30.6730	0.0117	16.8191	0.1136	Sudan	30.1523	0.0138	18.3859	0.0694
Swaziland	16.2597	0.4719	12.2517	0.3921	Swaziland	9.9421	0.9274	7.5162	0.8625
Togo	11.8782	0.8203	6.8169	0.9128	Togo	12.4016	0.7834	6.8868	0.9083
Uganda	20.5914	0.1974	15.2865	0.1786	Uganda	14.2734	0.6350	11.1015	0.5029
Hypothesis of at most 2 cointegration relationship					**Hypothesis of at most 2 cointegration relationship**				
Algeria	6.5889	0.3894	6.5889	0.3894	Algeria	5.4483	0.5333	5.4483	0.5333
Benin	6.7814	0.3679	6.7814	0.3679	Benin	4.0387	0.7365	4.0387	0.7365
Botswana	7.6822	0.2787	7.6822	0.2787	Botswana	5.8014	0.4859	5.8014	0.4859
Burkina Faso	4.3262	0.6945	4.3262	0.6945	Burkina Faso	4.3421	0.6922	4.3421	0.6922
Burundi	3.6679	0.7895	3.6679	0.7895	Burundi	1.9551	0.9714	1.9551	0.9714
Central African Rebublic	2.8031	0.8990	2.8031	0.8990	Central African Rebublic	3.2422	0.8466	3.2422	0.8466

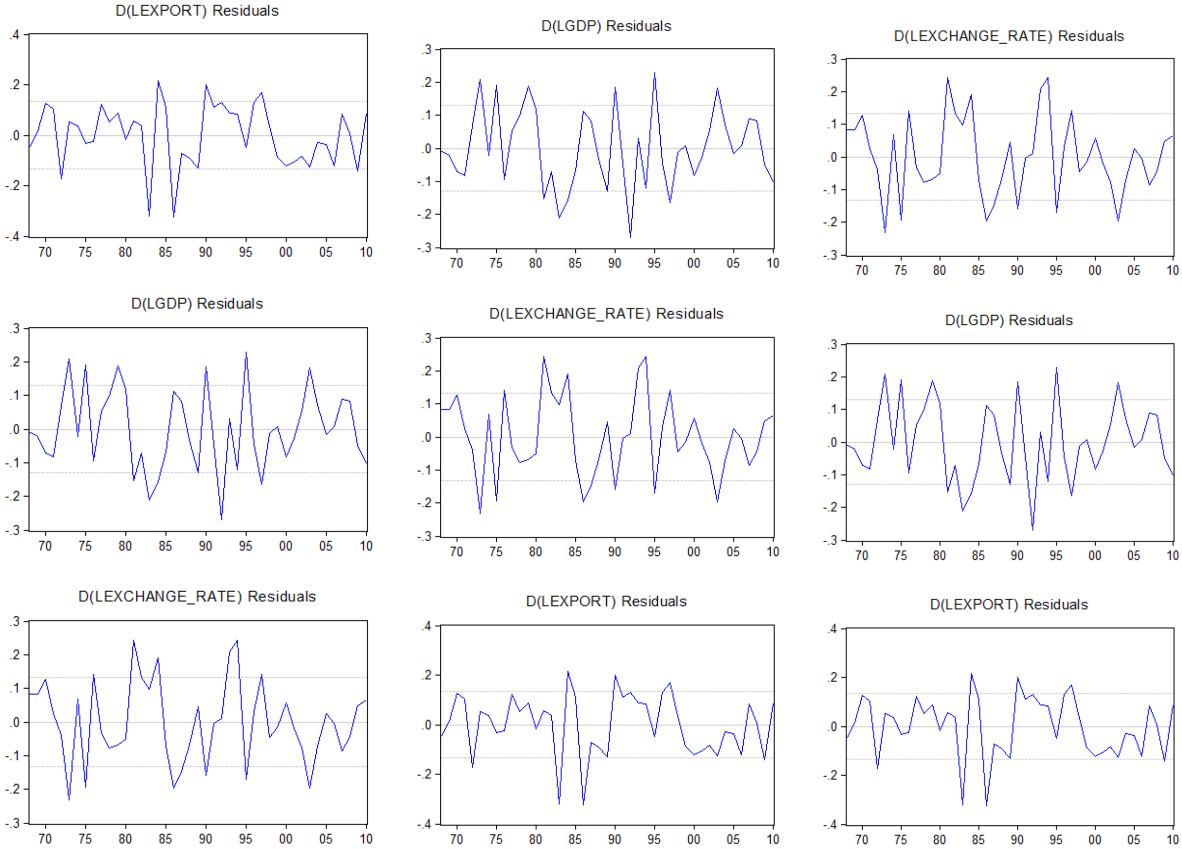
Cong, Republic	3.7405	0.7793	3.7405	0.7793
Cote d'Ivoire	5.2840	0.5561	5.2840	0.5561
Gabon	5.7055	0.4986	5.7055	0.4986
Ghana	2.9524	0.8822	2.9524	0.8822
Kenya	4.1292	0.7233	4.1292	0.7233
Lesotho	7.7382	0.2738	7.7382	0.2738
Cameroon	7.0363	0.3408	7.0363	0.3408
Madagascar	5.9055	0.4724	5.9055	0.4724
Malawi	4.3871	0.6856	4.3871	0.6856
Mauritania	3.2171	0.8498	3.2171	0.8498
Morocco	5.8648	0.4777	5.8648	0.4777
Niger	7.2453	0.3196	7.2453	0.3196
Nigeria	8.4772	0.2151	8.4772	0.2151
Rwanda	7.6029	0.2858	7.6029	0.2858
Senegal	8.0295	0.2493	8.0295	0.2493
Egypt	4.1687	0.7176	4.1687	0.7176
South Africa	5.8698	0.4770	5.8698	0.4770
Sudan	13.8539	0.0297	13.8539	0.0297
Swaziland	4.0080	0.7409	4.0080	0.7409
Togo	5.0614	0.5876	5.0614	0.5876
Uganda	5.3049	0.5532	5.3049	0.5532

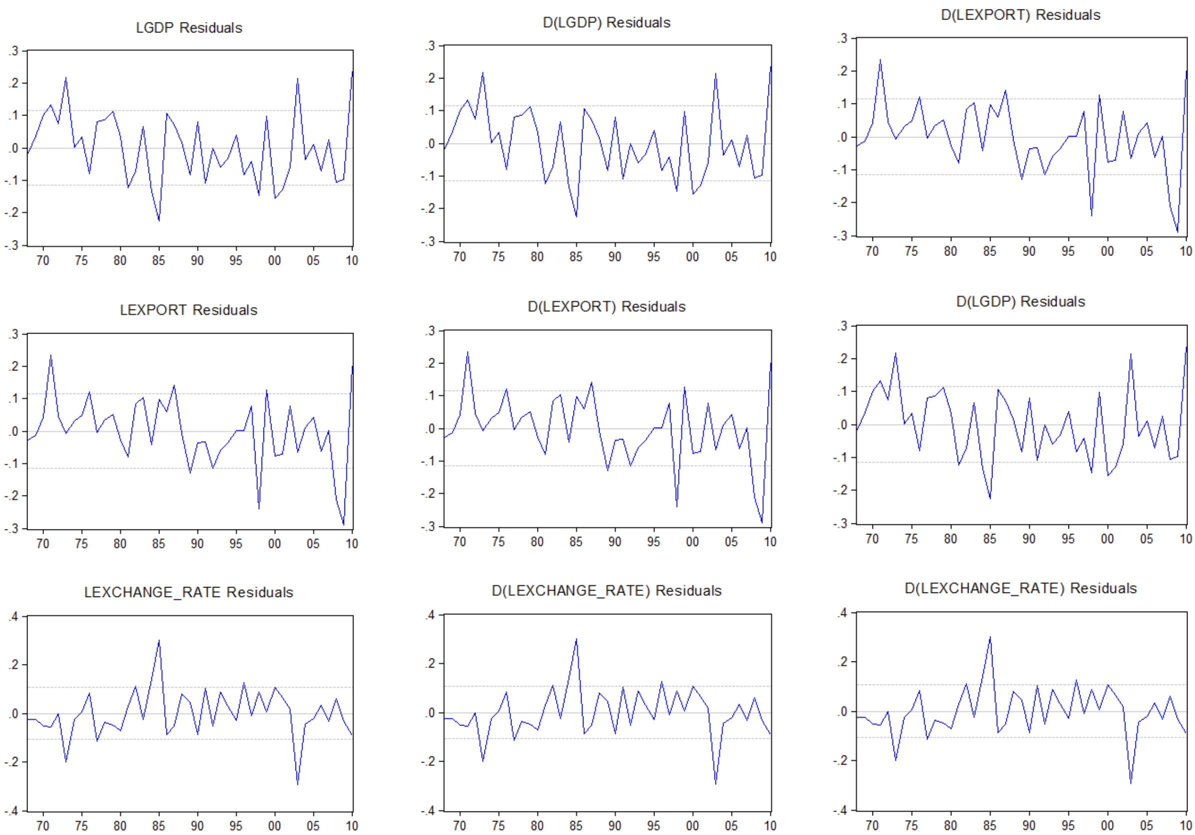
****MacKinnon-Haug-Michelis (1999) p-values**

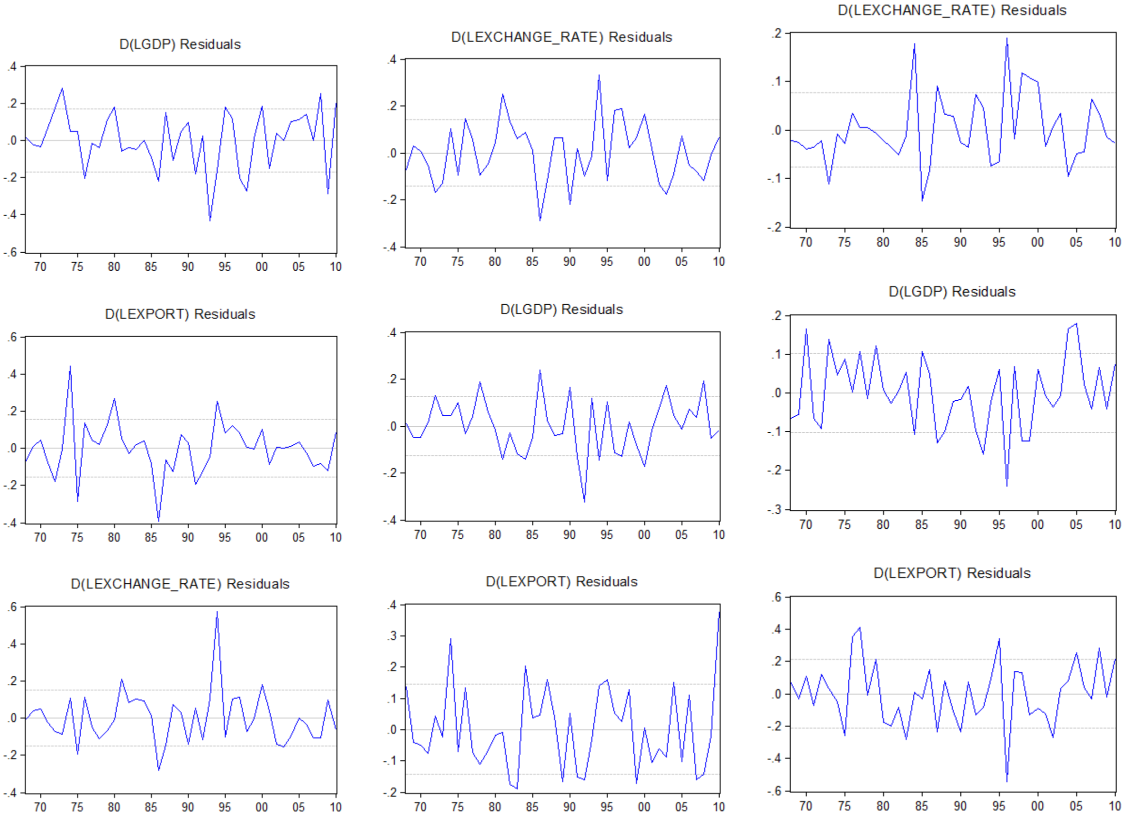
Cong, Republic	4.0512	0.7347	4.0512	0.7347
Cote d'Ivoire	6.2652	0.4274	6.2652	0.4274
Gabon	6.0439	0.4547	6.0439	0.4547
Ghana	2.9971	0.8769	2.9971	0.8769
Kenya	4.1478	0.7206	4.1478	0.7206
Lesotho	5.1095	0.5807	5.1095	0.5807
Cameroon	5.3882	0.5416	5.3882	0.5416
Madagascar	6.4881	0.4009	6.4881	0.4009
Malawi	5.0049	0.5957	5.0049	0.5957
Mauritania	0.8872	0.9996	0.8872	0.9996
Morocco	4.4637	0.6744	4.4637	0.6744
Niger	5.1922	0.5690	5.1922	0.5690
Nigeria	5.0379	0.5909	5.0379	0.5909
Rwanda	6.4516	0.4052	6.4516	0.4052
Senegal	6.2407	0.4303	6.2407	0.4303
Egypt	6.1218	0.4450	6.1218	0.4450
South Africa	3.3750	0.8293	3.3750	0.8293
Sudan	11.7664	0.0665	11.7664	0.0665
Swaziland	2.4260	0.9366	2.4260	0.9366
Togo	5.5148	0.5242	5.5148	0.5242
Uganda	3.1719	0.8555	3.1719	0.8555

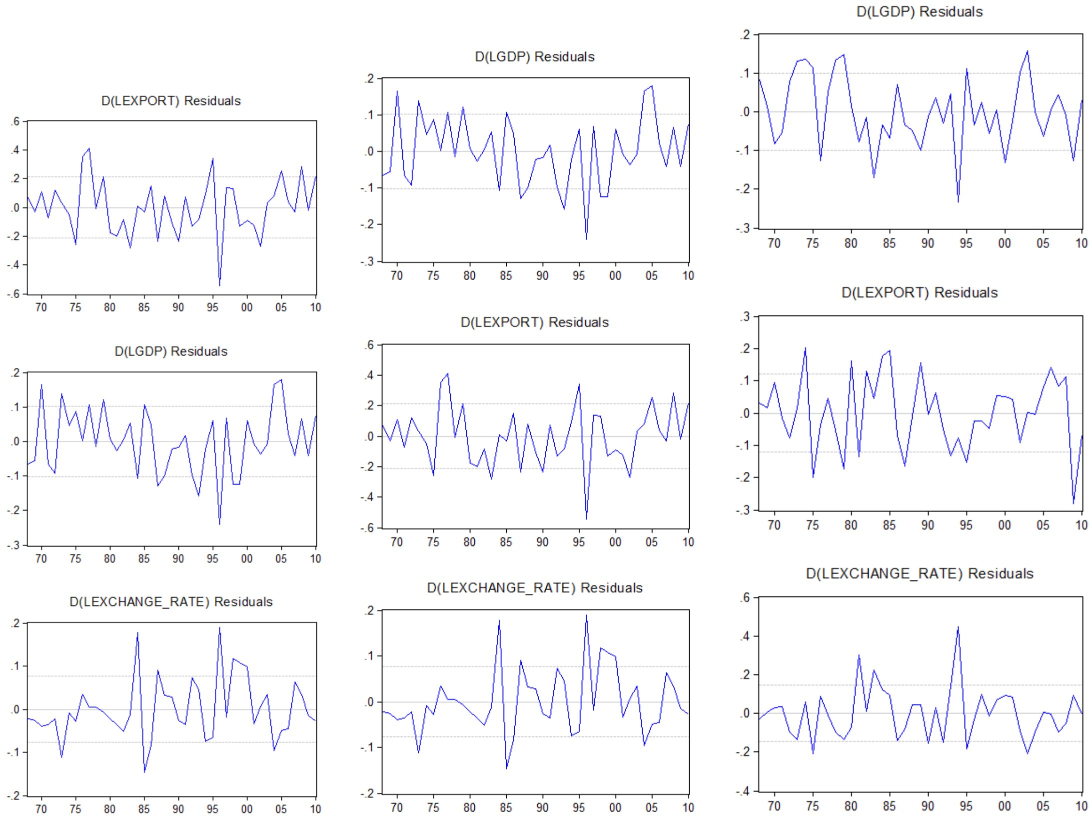
**MacKinnon-Haug-Michelis (1999) p-values

.2 Appendix C: Residual plots of the variables


D(LEXPORT) Residuals
D(LGDP) Residuals
D(LEXCHANGE_RATE) Residuals
D(LGDP) Residuals
D(LEXCHANGE_RATE) Residuals
D(LGDP) Residuals
D(LEXCHANGE_RATE) Residuals
D(LEXPORT) Residuals
D(LEXPORT) Residuals
70 75 80 85 90 95 00 05 10


LGDP Residuals
D(LGDP) Residuals
D(LEXPORT) Residuals
LEXPORT Residuals
D(LEXPORT) Residuals
D(LGDP) Residuals
LEXCHANGE_RATE Residuals
D(LEXCHANGE_RATE) Residuals
D(LEXCHANGE_RATE) Residuals
70 75 80 85 90 95 00 05 10





Residual Plot for the variables