

**FEMALE LABOUR FORCE PARTICIPATION AND ECONOMIC GROWTH IN SOUTH  
AFRICA**

**By**

**Sibongayena Sithembile Mabuyakhulu**

**Student No: 201065288**

**University of Zululand**



Submitted to the Faculty of Commerce, Administration and Law in Fulfilment of the  
Requirements for the Master of Commerce (Economics) Degree

**Female Labour Force Participation and Economic Growth in  
South Africa**

**By**

**Sibongayena Sithembile Mabuyakhulu**

**Student No. 201065288**

A Dissertation Submitted to the Faculty of Commerce, Administration and Law in  
Fulfilment of the Requirement for the Master of Commerce (Economics) Degree.

**University of Zululand**

**2019**

Supervisor: Prof I Kaseeram

The financial assistance of National Research Foundation toward this research is hereby  
acknowledged. Opinions expressed and conclusions arrived at, are those of the author.

## **DECLARATION**

I, Sibongayena Sithembile Mabuyakhulu, hereby declare that this dissertation saves for supervisory guidance received, is my own work and effort except where specifies. I have, to the best of my knowledge and belief, acknowledged all the resources of information in line with normal academic agreements. I further certify that the dissertation is original, and has not been submitted before at this or any other university for the award of any degree at any other university with the aim of obtaining a degree.

Sibongayena Sithembile Mabuyakhulu

Signature .....

Date .....

## **ABSTRACT**

Since the mid-90s female work compel investment has seen a stark increment (by 38%) boosting the general business levels. However, by global norms female labour force participation stays low and it is lower than for men averaging to a gap. Women's work remains characterised by domestic and cultural divisions. This study establishes the long-run relationship between female labour force participation and economic growth and unemployment in South Africa by using time-series data collected from various data source for the period of 1980 to 2015 (yearly) and 2008 to 2016 (quarterly). Empirical studies form both developed and developing countries indicate different results and also indicating a U-shape relationship between female labour force participation and economic growth. This study adopted the Cointegration Vector Autoregressive and Vector Error Correction Model (multivariate equations) together with cointegration equations (FM-Ordinary Least Squares, Canonical Cointegrating Regression and Dynamic Ordinary Least Squares) to establish the long-run and short-run relationships and the effect of economic growth and unemployment on the participation of women in the labour force. The estimate of yearly data shows that unemployment is positive and significantly influences the participation of women in the labour force in the long run. Economic growth exhibits the n-shape relationship with female labour force participation in the long run, hence, it indicates the opposite of what other researchers have found. The Vector Error Correction Model indicate insignificant effect of economic growth and unemployment on FLFP.

## **ACKNOWLEDGEMENTS**

Firstly, I like to thank my Almighty God for seeing me through this incredible journey, for being there and helping me through all the months of hard work until I completed my study.

I would like to express my sincere and utmost gratitude to my supervisor, Prof. Contogiannis, for his time and guidance throughout the learning process of this dissertation. Also, my deepest thanks go to Prof. Kaseeram: his commitment and great patience is highly appreciated.

My special thanks goes to my classmate, family and friends I just want to express my gratitude for support and comments they showed and made toward my work it really made my research more important and also progressive. I learnt a lot through the recommendations they made. It wasn't easy but I really appreciate their support and the patience they showed.

## TABLE OF CONTENTS

DECLARATION.....	i
ABSTRACT .....	ii
ACKNOWLEDGEMENTS .....	.iii
CHAPTER 1: INTRODUCTION .....	1
3.1.    BACKGROUND OF THE STUDY .....	1
3.2.    PROBLEM STATEMENT .....	3
3.3.    OBJECTIVE OF THE STUDY .....	3
3.4.    INTENDED CONTRIBUTION TO THE BODY OF KNOWLEDGE .....	4
3.5.    RESEARCH HYPOTHESES.....	4
3.6.    ORGANISATION OF THE STUDY .....	4
3.7.    ETHICAL CONSIDERATIONS.....	5
CHAPTER 2: THEORETICAL AND CONCEPTUAL FRAMEWORK .....	6
2.1.    INTRODUCTION.....	6
2.2.    LITERATURE REVIEW.....	6
2.2.1.    Theoretical Analysis .....	7
2.2.2.    Conceptual Literature.....	19
2.3.    CHAPTER SUMMARY .....	22
CHAPTER 3: SOUTH AFRICAN LABOUR MARKET AND EMPIRICAL LITERATURE.....	24
3.8.    INTRODUCTION.....	24
3.9.    THE SOUTH AFRICAN LABOUR MARKET.....	24
3.9.1.    The Labour Force: The Level of Unemployment and Employment in RSA .....	25
3.10.    EMPIRICAL LITERATURE .....	26
3.10.1.    South African Literature .....	27
3.10.2.    Empirical Evidence from Abroad.....	29
3.11.    SUMMARY OF EMPIRICAL EVIDENCE .....	36
3.12.    CHAPTER SUMMARY .....	39
CHAPTER 4: STATISTICAL ESTIMATION METHODOLOGY .....	40
4.1.    INTRODUCTION.....	40
4.2.    DATA ISSUES.....	40
4.2.1.    Data Sources and Sample Size .....	41
4.2.2.    Description of Variables .....	41
4.2.3.    Justification of Variable Mix and Sample Size .....	42
4.2.4.    Data Transformation and Descriptive Analysis .....	42

4.3.	SPURIOUS REGRESSION CONCEPT .....	43
4.3.1.	Stationarity of Time-Series .....	44
4.4.	UNIT-ROOT TEST .....	44
4.4.1.	The Dickey-Fuller Test .....	45
4.4.2.	Augmented-Dickey-Fuller Test (ADF) .....	46
4.4.3.	The Phillips-Perron Test.....	47
4.5.	LAG LENGTH SELECTION CRITERIA .....	47
4.6.	Cointegration Analysis .....	48
4.6.1.	The Engle-Granger Approach .....	50
4.7.	JOHANSEN'S MAXIMUM LIKELIHOOD COINTEGRATION TEST .....	51
4.8.	THE VAR/VECM APPROACH .....	53
4.9.	POST-ESTIMATION DIAGNOSTIC ANALYSIS .....	55
4.9.1.	Heteroscedasticity .....	55
4.9.2.	Autocorrelation .....	56
4.9.3.	Normality .....	57
4.9.4.	Multicollinearity .....	58
4.9.5.	Functional form.....	59
4.10.	CAUSALITY TEST .....	59
4.10.1.	The Granger Causality Test .....	60
4.10.2.	The Sims Causality Test .....	61
4.11.	MODEL STABILITY.....	62
4.11.1.	Chow Test .....	62
1.1.1.	CUSUM and CUSUM-Sq Test .....	63
4.12.	IRF AND FORECAST ERROR-VARIANCE DECOMPOSITION.....	64
4.13.	SINGLE EQUATIONS.....	65
4.13.1.	Fully-Modified Ordinary Least Squares (FM-OLS).....	65
4.13.2.	Canonical Cointegrating Regression (CCR) .....	66
4.13.3	Dynamic Ordinary Least Squares Test (DOLS).....	66
4.14.	CHAPTER SUMMARY .....	67
CHAPTER 5: EMPIRICAL ANALYSIS AND INTERPRETATION OF RESULTS.....		68
5.1.	INTRODUCTION.....	68
5.2.	MODEL SPECIFICATION .....	68
5.3.	DESCRIPTION OF DATA .....	69
5.3.1.	Graphical Analysis of Data.....	69
5.3.2.	Descriptive Statistics .....	71
5.3.3.	Correlation Matrix .....	73

5.4.	STATIONARITY TESTING .....	73
5.5.	THE VAR AND VECM ESTIMATION PROCESS.....	76
5.5.1.	Lag Length Selection .....	76
5.5.2.	Stability of the VAR .....	76
5.5.3.	VAR Estimates .....	77
5.5.4.	VECM Estimates .....	81
5.5.4.6.	Granger Causality .....	87
5.6.	SINGLE EQUATIONS ESTIMATIONS .....	88
5.6.1.	Fully-Modified Ordinary Least Squares (FMOLS).....	90
5.6.2.	Canonical Cointegration Regression (CCR) .....	91
5.6.3.	Dynamic Ordinary Least Squares .....	92
5.7.	GRANGER CAUSALITY TEST (1980 – 2015 TIME-SERIES).....	93
5.8.	CHAPTER SUMMARY.....	94
CHAPTER 6: SUMMARY, POLICY RECOMMENDATIONS AND LIMITATIONS .....		95
6.1.	INTRODUCTION.....	95
6.2.	SUMMARY OF THE STUDY .....	95
6.3.	POLICY IMPLICATIONS.....	97
6.4.	LIMITATIONS AND RECOMMENDATIONS OF THE STUDY .....	97
REFERENCES.....		99
APPENDIX B: SINGLE EQUATIONS RESULTS .....		108
APPENDIX A: MULTIVARIATE EQUATIONS FIGURES AND TABLES .....		108

## CHAPTER 1: INTRODUCTION

### 3.1. BACKGROUND OF THE STUDY

Computable information indicates an expansion in the participation of women in the labour force, moreover, indicating the most significant change in the supply of labour in South Africa in the previous couple of years. Expanded work efficiency has boosted the development of the economy, in line with various growth theories. A large increase in women's participation in economic activities was noticed in 1999 than in previous years, but that indicated that many women were searching for work rather than finding it, and most unemployed and independently employed women were generally participating in the casual area with low paying jobs (Posel & Casele, 2002).

Since the mid-90s the participation of woman in the labour force has seen a sharp increase by 38% advancing the overall employment levels. However, by international standards, where women's economic participation is at 47%, of women's participation in the South African economy remains low and it is lower than for men by about 14%. Almost 97% of women work either in traditional female occupations (domestic workers) or in the domestic and agricultural sectors. They are concentrated within positions which pay less and have high rates of turnover.

Historically, women have found themselves exposed to a kind of discrimination embedded in policies and patriarchal attitudes, which discouraged them from participating in the labour market. South Africa is no exception although different programmes and policies have been adopted to ensure fair and equal access to the labour market for woman nationally. Such interventions are not just important from the perspective of protection of women's interests, but many family units are absolutely or partially reliant on women's earnings.

The participation of women in the labour market has changed significantly nationwide, reflecting differences in the level of unemployment, economic development, levels of education and social norms. Female Labour Force Participation (FLFP) is impacted by a range of factors. One measurement that has been generally applied is the U-curve association between FLFP and economic development, which suggests that, after a threshold level of female participation has been attained in the economy, further absorption of women into the economy has a significant positive nonlinear impact on

economic growth. As indicated by Verick (2014), concentrating on estimating these threshold and exponential growth effects is critical for policy makers since the consensus view among many development economists is that FLFP is a driver of development, and these investigations into the labour force assist in quantifying a nation's capacity to develop rapidly. In many emerging economies, female employment fluctuates in response to economic shocks that adversely impact on many households, especially since women are largely employed in low value or labour-intensive jobs which are the first to go during a recession. Hence, a study of the behaviour of FLFP rates over a business cycle is necessary to better understand these dynamics in emerging economies like South Africa, for example.

In the literature, female work supply is seen as both a driver and a result of growth. As more women participate in the labour force, the economy can grow faster in response to higher work inputs. In the meantime, as the country develops, women's abilities also increase, while social rules weaken, empowering women to participate in work outside the home. Thus, policymakers need to evaluate the extent to which women's supply of labour in the market fluctuates and to monitor their participation in the labour force, as well as in the variety of vocations.

Economic theory indicates an indirect relationship between FLFP and economic growth. A rise in female economic activity has been due to the level of economic growth accompanied by the level of employment; however, empirical research is needed to investigate the existence of the U-shaped relationship between FLFP and economic growth. Theoretical and empirical (worldwide) evidence concentrates on this relationship, and found that the relationship is robust overtime and relatively stable. Several studies found that the U-shaped relationship holds (Olivetti, 2013; Tam, 2011; Tsani, Paroussos, Fragiadakis, Charalambidis & Capros, 2012). Most South African studies are qualitative and have been conducted using cross-sectional data. Some of the studies employed are purely descriptive analysis which does not indicate a clear relationship among the selected economic variables. Hence, this study employs quantitative analysis and time-series data to provide new information about the existence of the relationship (Festus, Kasongo, Moses & Yu, 2016; Ntuli, 2007; Posel & Casele:2002; Yakubu, 2010).

This study seeks to examine the long-run causality relationship between FLFP and economic growth in South Africa and determine the validity of the U-shape relationship. Furthermore, it investigates whether economic growth contributes to the level of women participating in labour force using time-series data from 1980 to 2015.

### **3.2. PROBLEM STATEMENT**

Over the past decades, FLFP in South Africa has been low in all sectors of the economy. However growth in the agricultural and service sectors has increased the number of women participating in economic activities (Yakubu, 2010). A number of studies have been conducted in various developed and emerging countries pertaining to an increase of the number of women in the labour market. In most emerging economies, a rising trend in female employment has been documented (ILO, 2005). In the South African context, only a few studies (Posel & Casele, 2002; Ntuli, 2007; Van der Westhuizen et al., 2007; Yakubu, 2010) have been conducted compared to other emerging countries and most of those studies have been qualitative studies using cross-sectional data. They tend to focus on short periods and quarters of the year which does not assist in evaluating and pin-pointing the main factors that influence the level of FLFP in the economy. The study by Posel and Casele (2002) which investigated the feminisation of the labour force in SA was inconclusive because of unavailability of the data; hence, their results are inconclusive.

This study seeks to establish macroeconomic determinants of FLFP in South Africa, based on quantitative analyses using recent time-series data.

### **3.3. OBJECTIVE OF THE STUDY**

This study seeks:

- To establish if FLFP, economic growth and unemployment are cointegrated.
- To examine the direction of causality between FLFP and unemployment.
- To examine the direction of causality between FLFP and output growth.
- To investigate whether the U-hypothesis can explain the relationship between FLFP rate and output.
- To look at the short-run dynamics between FLFP, economic growth and unemployment.

### **3.4. INTENDED CONTRIBUTION TO THE BODY OF KNOWLEDGE**

To the best of the researcher's knowledge, no study has been conducted in SA which investigated macroeconomics factors (gross domestic product [GDP] and unemployment) as the determinants of FLFP. By using time-series data and a large sample size and adopting a more sophisticated VAR-VEC model, the study aims to contribute to the existing debate by introducing and analysing the economy's GDP per capita and unemployment as major determinants of the participation of women in labour force. By doing so, it could assist policy makers to understand whether a high or low level of the mentioned variables increases or lowers FLFP in SA compared to other emerging economies. None of the South African studies has used time-series data; therefore, this paper intends to study the FLFP and economic growth relationship using recent time-series data.

### **3.5. RESEARCH HYPOTHESES**

- ❖ Hypothesis 1: Women's participation in the labour force, economic growth and unemployment are cointegrated.
- ❖ Hypothesis 2: The unemployment rate and output growth causes the female participation rate to rise in South Africa.
- ❖ Hypothesis 3: There is bidirectional causality between FLFP and economic growth in South Africa
- ❖ Hypothesis 4: The U-shaped hypothesis is valid for South Africa.
- ❖ Hypothesis 5: There are significant short-run dynamics between FLFP, economic growth and unemployment.

### **3.6. ORGANISATION OF THE STUDY**

This study consists of six chapters. Chapter 1 is provides an overview of the entire research paper. Chapter 2 discusses the main theories used in modeling the linkages between FLFP, economic growth and the rate of unemployment. The theoretical basis of women's participation in the labour force is rooted in the labour supply theories. Economists believe that improvements in human capital, a change in social norms and the drive for equal distribution of income can be associated with a faster pace of FLFP in the labour market. Moreover, the chapter gives more information on the conceptual literature on FLFP links to economic growth.

Chapter 3 discusses empirical studies on the relationship between FLFP and its determinants. Many existing South African and internationally empirical studies are reviewed in an attempt to identify general findings on the theme and identify gaps that this study may address. The empirical review helps in identifying appropriate variables and a creating a model to determine the relationship among selected variables in the South African context.

Chapter 4 covers the relevant statistical concepts while econometrics models are addressed in chapter 5. Sections 2, 3 and 4 of the chapter cover the time-series estimation methodology. Data issues, the concept of stationarity, cointegration and their respective test are discussed followed by the lag selection and cointegration in sections 5 and 6. Sections 7 and 8 discuss the VAR/VECM modeling technique and post diagnostic analysis. Causality and model stability are presented on sections 9 and 10. Lastly, sections 11 and 12 provide IRF/EVD and single equations estimation respectively.

Chapter 5 gives detailed information on the empirical estimation technique discussed in chapter 4, and presents and interprets the estimated model. As outlined above, one of the objectives of the study is to examine the long-term relationship between FLFP and its determinants and exploit this relationship to derive the short-run model, which is the VECM. This section also shows the outcomes from the single cointegrating models. The last chapter is a summary of the study discussing the overall importance of the study and making recommendations for macroeconomic policy. Moreover, strengths and shortcomings of the study which future researchers should address are mentioned.

### **3.7. ETHICAL CONSIDERATIONS**

The researcher takes into account peripheral issues governing good conduct in research. Such issues include collection of data and transformation utilising econometrics techniques. The researcher is aware of plagiarism rules. Correct citations are done to avoid plagiarism. The Harvard style of referencing is adopted to ensure correct referencing. Plagiarism detection software is used and the thesis was corrected where needed before submission.

## **CHAPTER 2: THEORETICAL AND CONCEPTUAL FRAMEWORK**

### **2.1. INTRODUCTION**

This chapter of the study focuses on the theoretical framework underpinning the levels of labour supply from a broad perspective. The aim of this chapter is to review the literature of different macro- and microeconomic theories on the participation of women in the labour force in South Africa. It first provides an overview of the theoretical links between labour supply or labour force and economic growth, followed by the conceptual literature which also addresses the relationship.

Detailed discussions of different theoretical views explain different approaches concerning the labour force. The role of economic growth in promoting participation in the labour force is discussed in detail, linking it to the theoretical analysis. This part of the study highlights basic theories that have been used to support the relationship between FLFP and economic growth. In line with the definition of FLFP in Chapter 1, this chapter presents a formal discussion of theories of the labour supply in general.

In this chapter, various theoretical and conceptual issues are debated based on the participation of female in the labour force, particularly in the South African economy. The theoretical analysis includes economic theories that explain the nature of the participation of women in the labour market. These theories also provide a justification of the determinants of female labour supply in the economy. Other than theoretical exposition, the conceptual argument of macroeconomics explains some determinants of FLFP. There is no direct theory of macroeconomics that explains the participation of women in the labour market, hence FLFP is a conceptual issue.

Basically, this chapter is categorised into two broad sections which include the theoretical literature explaining labour supply and growth of the economy in the first section, followed by the conceptual analysis.

### **2.2. LITERATURE REVIEW**

This section of the study provides the literature review on FLFP rate and economic growth; it is divided into two parts. The first part discusses the main theories of labour supply and economic growth while the second part provides the conceptual framework discussing the relationship between the female participation in the labour market and

economic growth. Notably, several theories hypothesise a U-shaped relationship between FLFP and growth.

### **2.2.1. Theoretical Analysis**

Economic theory indicates that supply and demand are the basic concepts that assist in understanding the market economy, but the main focus of this paper is supply and demand in the labour market. The labour demand curve is downwards sloping due to the fact that there is less demand for labour when the wage rate is high, while the labour supply curve is upward sloping because, at a high wage, the opportunity cost of not working is high, with people being more willing to supply labour. The following theories discuss how and when individuals supply labour as one of the factors of production in the market.

#### **2.2.1.1. The theory of the decision to work**

The theory explains how individuals make decisions on how to spend their lives. Individuals spend their lives in gaining employment, taking decisions on how much to supply in the labour market, or when they should supply this labour, and which position or industry to specialise in (Sloane, 2013). People spend their time in different ways; for example, one may decide to use his/her time to work for pay, spend it on leisure activities, or do household chores which do not involve any payment. The assumption which states that work for pay and work in the household is different is not considered in this theory; both work for pay and household work are treated as working for pay. The decision to work is basically a choice between working for pay or leisure.

The analysis of labour/leisure choice is analysed by looking at the factors that affect the demand for a good. The demand for a good is a function of opportunity cost, wealth, and preferences.

$$D = f(\bar{C} W) \quad (2.1)$$

The above equation indicates the demand function of a good where  $C$  represent the opportunity cost and  $W$  is wealth for an individual, depending on preferences. The notation — above cost indicates the direction of the variable to the demand holding other things constant. Ehrenberg and Smith (1991) indicate that most economists assume that preferences are fixed and given at any point in time. To understand the

demand for leisure analysis we must first know the opportunity cost of leisure which is the cost of spending one-hour shopping as to what one could earn if he /she had spent that hour working. The opportunity cost of leisure is basically the wage rate.

Secondly, wealth must be measured according to family's belongings (physical assets, financial accounts and investments). Human skills are also considered as the part of the wealth or assets that is rented for a price. The larger the value of human assets, the more an individual can receive in wages. Individuals' total income is used to measure people's wealth since it is impossible to measure the wealth of an individual. If we replace the demand function in equation (2.1) by the demand for leisure function we obtain the following equation assuming that leisure is a normal good:

$$D = f(\bar{W}, \bar{Y}) \quad (2.2)$$

Where  $D$  denotes the demand for leisure,  $w$  is the wage rate and  $Y$  denotes income. The sign above the explanatory variables indicates the direction of leisure whenever there are changes in these variables, keeping other variables unchanged. The equation indicates that, *ceteris paribus*, if income is increasing, leisure demand also rises leading to the income effect which is the preferred hours of leisure to income changes, holding wages constant. The increase in income induces individuals to consume more leisure, holding leisure cost constant. Algebraically notation of the income effect is expressed as follows;

$$\text{Income effect} = \frac{\Delta H}{\Delta Y} \uparrow \bar{W}$$

The income effect is defined as a ratio of change in hours of work ( $\Delta H$ ) to a change in income ( $\Delta Y$ ) holding wages constant. An increase in wage rate reduces the demand for leisure, holding income constant and work incentive increases. Hours of work and leisure are substituted for each other, holding income constant and the substitution effect is created.

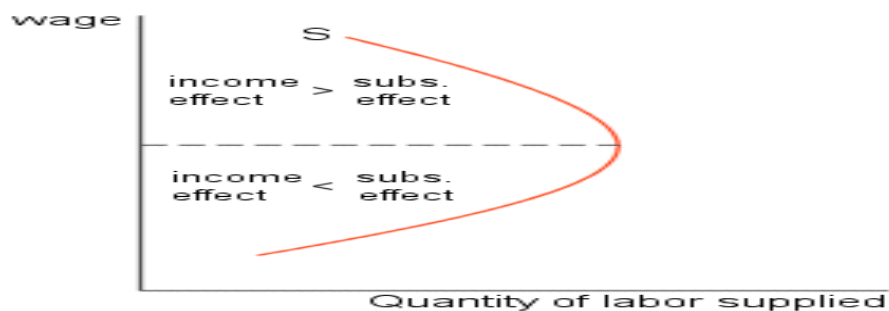
$$\text{Substitution effect} = \frac{\Delta H}{\Delta W} \uparrow \bar{Y}$$

Both income and substitution effect occurs due to a rise in the wage rate but works in the opposite direction. The income effect is negative while the substitution effect is positive to the supply of labour in the market. Due to the presence of both effects

working in the opposite direction, it is impossible to predict the overall labour supply in the labour market (Collins, Bosworth & Rodrick, 1996). The response of labour supply to changes in wages involves both effects.

Individuals decrease their labour supply when wages increase only if the income effect is dominant. Although the substitution effect may be present, it is not enough to discourage or to prevent the declining labour supply. However, it acts as a moderating influence. In a case where the substitution effect is dominant, workers increase their supply of labour due to a rise in wages and the supply curve is positively sloped, unlike in the domination of income effect where the supply curve is negatively sloped (Figure 2.1).

**Figure 2.1: The labour supply curve**



Source: (Varian, 2014)

People are willing to increase their hours of work when wages go up as long as wages are greater than the reservation wages where the substitution effect is greater than income effect. However, a further increase in wages results in declining hours of work indicating that income effect is now greater than the substitution effect.

### 2.2.1.2. The labour supply theory

The labour supply theory develops the idea of a worker as a rational decision maker, taking decisions (choices) based on various opportunities (bad and good) presented in the labour market. It also states how individuals earn a living by providing labour services to the market, while there are other important factors of production that bring changes in the market. The guiding theoretical question says 'is labour supply an increasing function of real wage rate?' As a worker's wage increases, does the supply of labour increase too?

This model is also known as income or labour-leisure model, and is applicable in terms of hours and participation. This study adopts this kind of model based on participation to analyse the effect of economic growth on the participation of labour in the market. The supply of labour depends on changes in the real wage. The following is the derivation of an individual labour supply curve indicating how income and substitution effects affect the supply of labour in the market (Bosworth, 1996).

The theory begins by deriving the labour supply curve in relation to the market wage. Economists have divided it into whether to work (participate) and how much to work (working hours) components; therefore, the model was developed to investigate how participation changes in a given economic environment (output). The participation of individuals can be explained in terms of substitution, income effects and in terms of reservation wage. Changes in the price of leisure relative to goods which follow a change in wage rate create the substitution effect in the labour market. Leisure is found to be expensive; from theoretical literature, this effect induces individuals to work longer hours as the wage rate increases. An increase in wages allows individuals to consume leisure at a high level and achieve a high level of satisfaction.

The participation decision of an individual in the labour force is more concerned about their willingness to work at the going wage rate. Most individuals participate in the labour market when their desired hours of work are greater than zero at any given wage rate. Individuals are willing to supply more hours of work as the offered wage rate increases. The increase in wages induces individuals to supply more labour to the market, but as it approaches a point where wages reach an upper limit which may be below the technical maximum of 100% of the participation rate, effort decreases. The supply of labour in the market works together with the effort that individuals are willing to provide in the workplace. It is noticeable that the intensity of effort or participation is expected to be affected by many factors which include the wage rate. Bosworth (1996) indicates that at any given real wage rate, a rise in wages motivates workers to improve their working pace, enables them to achieve a better standard of living and follow a good diet which also leads to increased productivity. The backward bending curve of the labour supply explains all the changes in income. It indicates that income increases and reaches a certain point (upper limit) where an increase in wages results in a reduction of effort in the market.

### 2.2.1.3. Reservation wage

The supply of labour theoretically involves the concept of the utility-maximisation framework. The simple neoclassical model under the supply of labour indicates that an individual's utility is maximised subject to their budgetary constraints. The supply of labour is based on the reservation wage which is the minimum wage that induces individuals to participate in the labour market. Individuals compare their reservation wage ( $w^r$ ) with the ongoing wage rate ( $w$ ) in the market (Bosworth, 1996).

The decision whether to participate in the market activity is based on whether the reservation wage is less or greater than the market wage. If the reservation wage is less than the ongoing wage, then the individuals will choose hours of work that maximise utility. The opposite is also true: if the reservation wage exceeds the market wage, then individuals may decide not to work and rather to enjoy leisure, hence hours of work equal zero.

$$W_r < W; H^* > 0$$

or

$$W_r > W; H^* = 0$$

Ntuli (2003) indicates that the reservation wage depends on the characteristics of an individual (race, age, marital status) and non-labour income whereas the ongoing wage depends on the skills or work experience that an individual has (human capital). Theoretical literature postulates that investing in human capital raises the opportunity cost of not working, and consequently increases the participation rate of individuals (Sloane, 2013). If the husband's income is high, women's reservation wage is also greater than the ongoing wage rate, and it lowers the chances of women to participate in market activities (Killingsworth & Heckman, 1986). When considering the reservation wage, the supply of labour is indicated by a U-shape between the age & FLFP and income & FLFP (Ntuli & Wittenberg, 2003).

### 2.2.1.4. The effect of changes in labour demand

Discussions on labour demand began in the 1930s and has changed in recent years with the idea of the added-worker hypothesis which states that when the household

breadwinner becomes jobless, another member of the family enters the labour force with the aim of maintaining family income and needs. However, Filer (1996) argues that the labour force participation should increase as unemployment is rising. The added-worker hypothesis indicates that during a recession, the level of unemployment is overestimated which means that the creation of a single job reduces unemployment by almost two or more individuals. On the other hand, the discouraged-worker hypothesis implies that job searching during periods of high levels of unemployment discourages people in a way that some unemployed individuals quit looking for a job and move out of the labour force.

The reduced demand for labour in the market has an impact on the individual choice about participation in the market; the aggregate demand reduces the family income because people who may lose their jobs or, in some cases, work fewer hours than before. The non-labour income (those who are not part of the labour force) reduces the reservation wage since income from other sources or family members diminishes and cannot satisfy all family needs. Family assets determine how many hours of work to supply in the market; for example, in a case where a family has a large pool of liquid assets or can easily lend at a given interest rate then the reduction on income earned by the participants in the market has an insignificant effect on participation in the labour force. A family that does not have any assets or has only a few assets and cannot borrow according to their wishes, finds interest rates on loans to be very high and unemployment benefits to be less than the ongoing wage and insufficient to cover all family needs (Filer, 1996). Therefore, for these conditions, an increase in labour force participation by one family member means an adjustment to employment.

The discouraged-worker effect pinpoints the importance of declining market opportunities even during the recovery stage of an economy, particularly for labour force participation. When labour demand falls, the wage rate falls, available jobs pay less and become less attractive than before, and it is sometimes not easy to find work, which is the reason why individuals decide to quit searching for a job. This effect favours nonmarket activities for individuals trying to make decisions based on schooling, training or doing home activities rather than looking for employment. When the unemployment rate is very high (low employment), workers who lose their jobs may decide to retire early before their due retirement age than during the time of low

unemployment (high employment) because finding a suitable position is easy at a low level of unemployment.

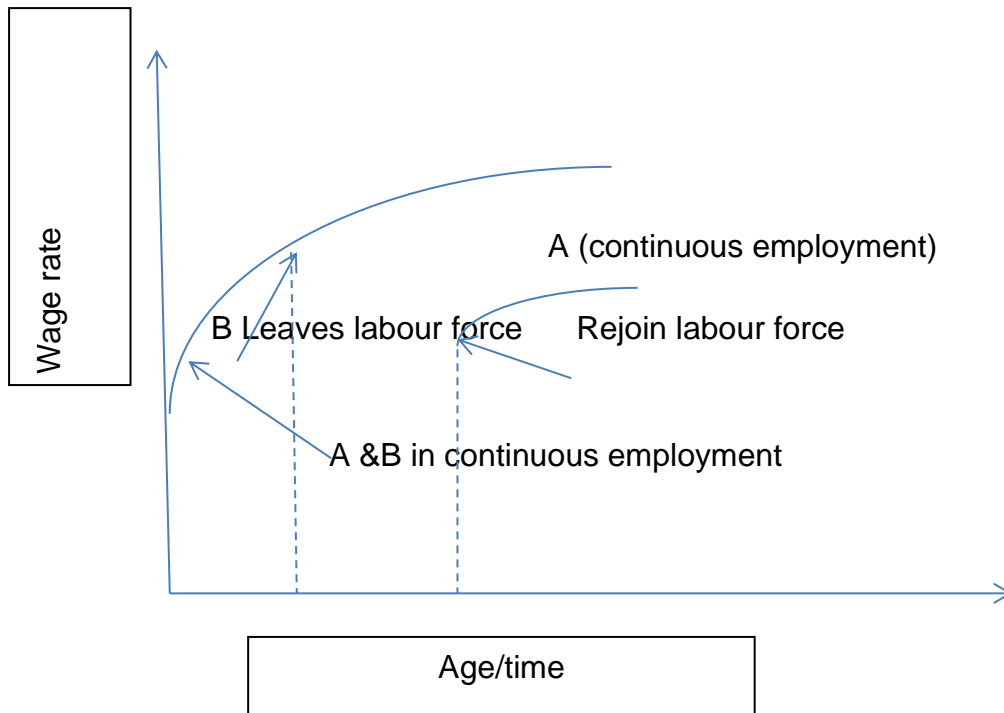
The added-worker and discouraged-worker effects can be present at the same time for different individuals or household. Economic theory underpins that women whose husbands cannot earn as much or work as many hours as they wish are more likely to be part of the labour force but, due to the low rate of employment and difficulties in finding work, individuals may be discouraged from looking for work and some may decide to leave the labour force.

#### **2.2.1.5. The lifecycle model**

The model focuses on labour supply decision based on the hours of work (job choice) and participation in a lifecycle or a dynamic framework. Dynamic labour supply models include models where real wages are either exogenous or endogenous. Basically, the primary goal of the model is to explain how much labour an individual supplies in the current period rather than in future, and thereby to estimate the total value of the supply of labour over a lifetime.

Killingsworth (1983) indicates that the behaviour during the lifecycle of both women and men remains less or stable, even though the pattern indicates changes in recent years. In this model, wages are classified as endogenous and exogenous. An individual's labour supply decision is constrained by the skills possessed in the short run but in the long run, they are able to adjust their capacity to work or skills through investing in training or in education. Killingsworth (1983) studied the behaviour of men and women in a lifecycle using different kinds of data. The men's trend indicates that labour supply and wage rate are concave. Men spend their first part of life in training and full-time education. They work more in the first stage of entry in the market and less continuously up until retirement stage. Women did not follow the same pattern but moved in and out of the labour market after their education and training. The model explains that individuals go for training over their lifecycle because, in the early stages of life, the wage rate is low; they, therefore, forgo the earnings and undertake training. However, earnings increase over time and the cost of not working (training) increases as times goes on. The training value becomes negative and so no further training is undertaken by women (Bosworth, 1996).

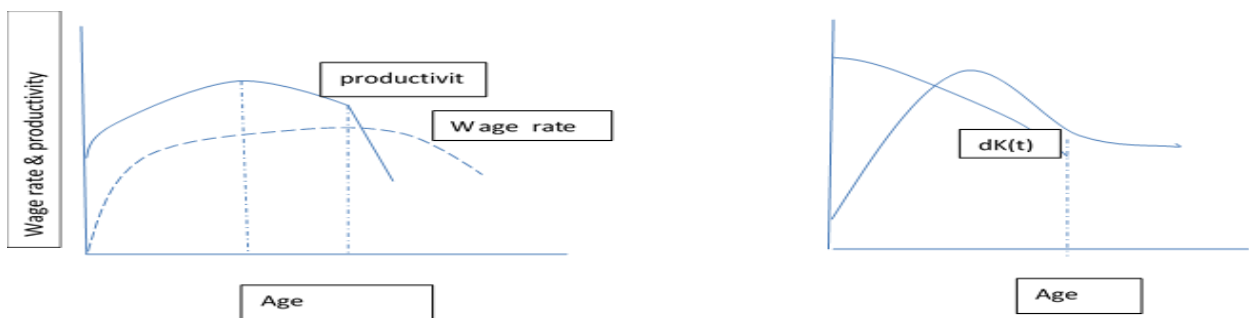
**FIGURE 2.1: WAGE PROFILE**



Source: Bosworth (1996)

The above diagram indicates that women move in and out of the labour force after leaving school. The average across women is a combination of the two individuals A and B because some women are productive less or more continuously up until the retirement age while some move in and out of the labour force or and others do not spend their lifetime in the labour force.

**FIGURE 2.2: (A) PRODUCTIVITY AND WAGE RATE OVER THE LIFE-CYCLE. (B) INVESTMENT IN AND DEPRECIATION OF HCT OVER THE LIFECYCLE.**



(a) Productivity and wage rate over the life-cycle

(b) Investment in and depreciation of HCT over the lifecycle.

The marginal product of human capital [ $K(t)$ ] and depreciation [ $\delta K(t)$ ] are the main determinants of the patterns of productivity in the market over time. Human capital depreciation increases over time although depreciation alone does not change (constant) as human capital is growing [ $K(t) > \delta K(t)$ ], thus human capital declines over time as the level of investment (training) falls, therefore  $\delta K(t) > K(t)$ . This indicates a reduction in the total value of individual's human capital. The implication of this model is that individuals' market productivity (potential wage) tracks the stock of human capital, and then the U-shape is created. There is a negative relationship between individuals' wages and investment in training but the wage rate is positively related to productivity (Bosworth, 1996).

The dynamic labour supply model with exogenous wage indicate that the assumption which states that individuals consume all of their income in every period are not considered. An individual's wealth, income accumulation, and expenditure determine the individual's budgetary constraints; thus, individuals have to maximise their lifetime utility given the budgetary constraint. The constrained maximisation yields important behavioural effects of an individual labour supply such as efficiency effect, which allow individuals to participate more in the workplace when the wage rate is high. The interest rate effect occurs when individuals work hard at first and less at a later stage (saving some of their earnings for the high rate of their compound interest in future). Lastly the time preference rate effect also induces individuals to work less at first and more at later a stage; basically, the opposite of the efficiency effect. Mathematically, the utility-maximisation decision is expressed as follows:

$$\dot{H}_t = a \frac{\dot{w}_t}{w_t} + b(\phi - r) \quad (2.3)$$

The first term of the right-hand side of the above equation indicates the efficiency effect while the second term is the time effects which is divided into two parts: the time preference rate effect denoted by  $\phi$  and interest rate effect denoted by  $r$ . The sign of the time effect depends on whether the interest rate is equal, greater than or less than the preference rate.

Models assuming that wages are endogenous only focus on education and training activities of an individual whereas models assuming that wages are an exogenous focus on the broader perspective. Fixed costs of seeking employment should be

considered when deciding to quit the job or rejoin the labour force. The opportunity of getting a new job as the market wage is greater than the reservation wage depends on the demand factors of the market (Barker, 2013).

Under the dynamic labour supply model with endogenous wage, the wage rate is a function of individuals' productivity in a workplace; hence productivity is determined by the stock of human capital. Wages are also influenced by individuals through investing in human capital. Individual's tradeoff between current and future earnings whereby they forgo part of their current income for higher wages in a future period, which introduces the idea of training effect. Individual's tradeoff between high wages in future generated from additional training or education and valuable leisure time they have forgone. Some individuals decide to spend their time in market work, which gives them more experience and increases their productivity leading to a high wage rate (Killingsworth, 1983).

The increase in wages (wage effect) induces individuals to increase their labour supply. Wages are found to depend on the nature of employment. The continuing depreciation of human capital explains the inability of women to compensate for investment in human capital because they are more involved in family activities and caring.

#### **2.2.1.6. Human capital theory**

The study adopts the marginal productivity analysis recently known as HCT. This theory helps to explain how to determine the wage rate in the market. This analysis assumes that wages in the labour market are determined by the knowledge and skills an individual possesses. The development of human capital began as part of a drive to enhance overall effective firm performance and minimise costs.

Human capital is defined as the skills and knowledge and other professional training that individuals have as reflected by their marginal productivity. It is the major factor which improves employees' capacity to work and is regarded as a firm's asset (Marimuthi et al, 2009). Rastogi (2000) emphasises that human capital is an input to the market especially where human effort is needed. Employees need to undergo certain training or schooling in order to obtain skills and knowledge which improve their participation in the labour market. Improvement in the business or market environment

enables firms to obtain a competitive advantage through their well-trained, educated and innovative employees. Investment in human capital is important for a firm's long-term sustainability. Human resource input plays a significant role in the firm to achieve competitiveness.

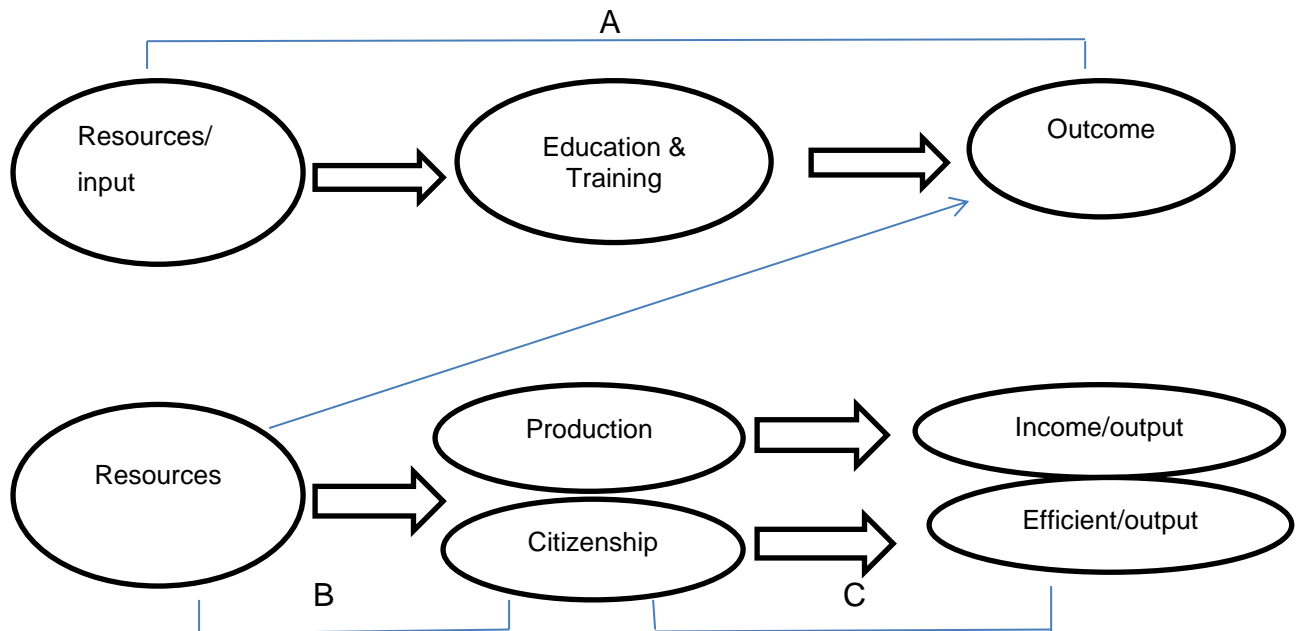
The classical economic view indicates that human capital assumes that labour is a tradable commodity in terms of scale (Marimuthu, Arokiasamy & Ismail, 2009). This classical theory emphasises the exploitation of labour by capital. Investing in human power is noted as the most important and valuable capital in the labour market. Human capital can be viewed from different perspectives, namely, economic and social (Becker, 1993). Economists distinguish between the two human capital hypotheses, the human skills that can be obtained and used in a specific firm or job (firm-specific human capital) and those skills and knowledge that can be applied across the market or industries (general human capital). Education and training are the most important and valuable human capital investment that individuals should consider when deciding to participate and how much labour to supply in the labour market. This kind of investment enables them to be more productive in the workplace.

The HCT focuses on organisations and individuals and is further discussed by Garavan Morley, Gunnigle and Collins (2001) that it contributes to (i) competitiveness of the organisation; (ii) job opportunity (indicating that it is easy for an individual to get a job if he/she is educated or trained); (iii) adaptability; and (iv) the individual's competitiveness in the market. This study focuses only on how individuals use their human power in the labour market and how much they supply at a given wage rate.

According to the HCT (Becker, 1986), people invest in the skills that will make them more productive and employable. This theory links directly to the central focus of this study by concentrating on how people supply labour according to the skills and knowledge they have. Human capital increases over the years as an individual gets a better education, training, and work experience. People search for jobs in order to escape from underemployment and unemployment. The higher the level of education of the unemployed person, the greater is the loss of income or return on the investment made in education, and the more intensive is the job search. It is believed that the shortages in the education and training system have an impact on labour supply and

economic growth. In the modern economy, work is increasingly knowledge-based, and those who are knowledge-enabled have a greater advantage in the labour market.

**Figure 2.3: A simple diagram indicating the relationship in the HCT and underlying assumptions**



Source: Marimuthu et al. (2009)

- Relationship A: indicates the production process that takes place when education and training are obtained. The underlying assumption based on this relationship is that investing in education and training results in individuals obtaining more information or knowledge about the production which also increases production.
- Relationship B: shows the relationship between increased learning and individual productivity. This relationship assumes that education raises the participation rate in the market which may be due to the high level of labour force participants.
- Relationship C: is the indication of the relationship between productivity, high wage rate, and business earnings. Productivity leads to high wage rate and high business earnings. The main idea of the above relationship is emphasises that human capital contributes to an increase in the participation rate in the labour force which also intends to raise income level received by both employees and employers.

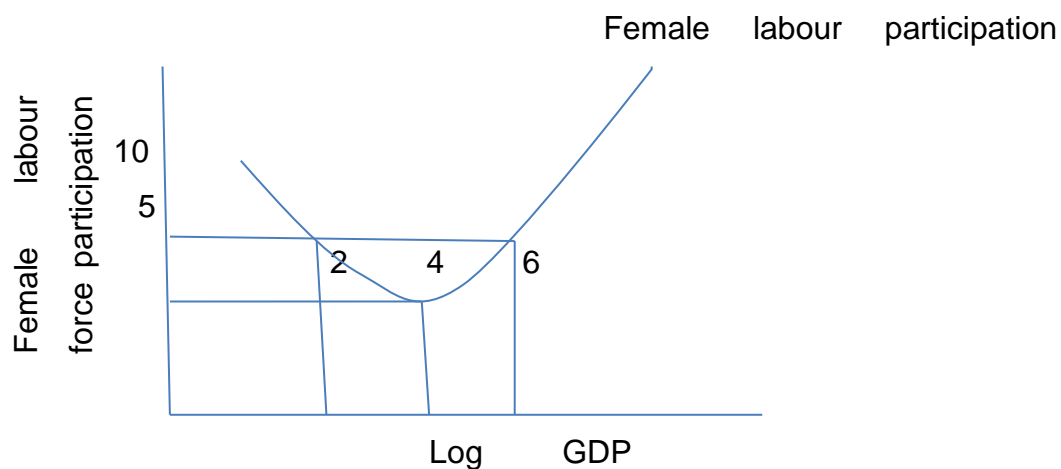
In this study, the understanding of human capital is based on how individuals supply their labour in the market once they have gained knowledge through investment in

education and training. Literature based on investing in skills (Garavan et al., 2001; Marimuthi, 2009) also indicates that human capital plays a significant role in creating a competitive advantage at an organisational level. A firm's human capital tends to lower production costs and raise production performance in the organisation.

## 2.2.2. Conceptual Literature

2.2.2.1 The U-shape hypothesis the U-shape hypothesis postulates that there is a trade-off between female access to the labour market during the early stages of economic growth. However, further increases in growth shift the economy from the industrial to the service sector. Low economic growth induces high women participation in the labour force. It is noticeable that with an increase in a country's output, women lower their participation to a certain level where after it begins to rise again (Forgha & Mbella, 2016).

**Figure 2.4: The U-Curve**



The high participation rate in the labour force during the low level of growth reflects the added-worker effect (Makepe & Oageng, 2012). The added work hypothesis indicates that when the household breadwinner becomes jobless, other family members join the labour market with the aim of maintaining family income.

Many theories have hypothesised that the U-curve relationship that exists between female labour participation and economic growth. Tam (2011) reveals that the income and substitution effects play a crucial role in illustrating the U-shape pattern between

the two economic variables. When income levels are low, the agricultural sector dominates in such a way that more women enter the labour force. As FLFP rise, income increases and family demands rise. However, Goldin (1995) argues that such an income effect results in a low level of women participating in the labour force. The substitution effects arise during the introduction of capital goods where different sectors become more capital-intensive. Women are substituted by machinery due to the advantage of male physical strength compared to women in the labour market. Both substitution and the income effect lower the level of FLFP; hence, as the level of development (growth) proceeds further, mental human capital becomes more valuable (through education) than physical human capital. During periods of rapid economic growth, real wages rise and the substitution effect encourages the participation of women in the labour force and overcomes the income effect which discourages FLFP; therefore, income will rise as female participation rises; thus, the U-shape is developed.

The downward portion of the U-curve explains movement from the agricultural sector to the industrial economy and where more male strength is required (Lechman & Kaur, 2015). During periods of low income levels, women bring their family members to participate on family farms, but as income rises, changes occur that reduce women's engagement in the labour force. The agricultural sectors become dominated by the manufacturing sector and economic development shifts from farm production to manufacturing and other paid work. The increase in production outside family farms leads to people moving from household productivity to paid work which contributes to a rise of the labour force.

The increasing rate of female participation in the labour market is explained by the increasing portion of the curve and indicates that the economy is service based (Olivetti, 2013). Economic development beyond low income and decreasing female participation are regarded as the results of income effect, the wage that the husband receives and unemployment benefits (non-labour income) which discourage women from participating in the market. Sackey (2005) postulates that high growth rate has a positive effect on female participation and high education attainment.

When new technology is adopted in production, women are displaced in the labour force (substitution effect) and, during the early stage of development, women gain

more access to education, which allows them to be more active and improve their participation level in the economy, which also creates a U-shape relationship between female participation and economic development.

A low level of income shifts female participation to family farms leading to a rise in the agricultural sector (Tam, 2010). The decreasing part of the U explains the fall in FLFP and economic development; as a result, the decrease in economic activities contributing to the higher level of real GDP leads to a fall in women's participation, and when the economy starts to grow, economic development rises with FLFP.

Market failure and implementation of restrictive policies in the labour market tend to discourage women from participating in the labour force. Jaumotte (2013) revealed that by removing such policies may stimulate a country's economic growth and women would be more economically active which improves their labour force participation rate.

The dynamic panel method has been used to identify the effect and relationship of education and economic development. The U-shape make it easy for economic policymakers to design rational policies to deal with the negative effects that contribute to a declining FLFP rate. Since growth in the agricultural and manufacturing sectors has been decreasing, the woman's economic activities have not contributed to the economic development programme, because the changes in these two sectors have a negative effect on the woman's participation rate.

#### **2.2.2.1. Labour force segmentation**

The theory of labour force segmentation by gender implies that there are certain jobs restricted to men only and others to women, but female wages are usually lower than those paid to men. Employment and unemployment surveys have revealed that most countries have been facing a decrease in FLFP rate but with an increase in economic growth and educational enrollment level for many years. This decline in FLFP has been assumed to be due to structural changes, income and substitution effects that affect women's activities in the economy. With an increase in capital, the economy grows and there is a structural shift from agriculture to industry and the services sector which also explains the changes in the female participation rate. The agricultural sector can employ fewer people, and women cannot move to other sectors for

employment immediately because of household activities (childcare), a lack of required skills and knowledge.

The structural shift in the economy has a negative effect on FLFP rate, especially in South Africa compared to other countries, and has shown that the agricultural sector has declined but not at the same level as the manufacturing sector while the service sector is enjoying a rapid growth rate in terms of value-add. Structural adjustment programmes, low economic growth and high poverty levels in South Africa have forced women to be more active in economic activities that lead to high level of development.

Women's employment is generally characterised by inferior conditions to those faced by men, meaning that they are employed under 'atypical employment' (temporary, casual or part-time job). Usually, when they enter the labour force, they have to work in different sectors from men.

The formalisation and lack of employment have made the labour force more complicated for women to participate in or contribute to economic activities. The employment rate has been low due to low economic growth and has not been seen to improve women's economic status. The interrelationship between net domestic product of some countries and women's economic activities has been studied using the U-shape hypothesis, which revealed that there is a decline in the FLFP rate. Although economic growth is increasing, women's participation will only increase when the economic growth rate stabilises.

### **2.3. CHAPTER SUMMARY**

The aim of this chapter has been to present the theoretical and conceptual literature underpinning this study. The main focus was to clarify the linkage that exists between FLFP and economic growth. The theories discussed indicate that the participation of women in the labour force can be determined at a macro- and micro-level. At a macro-level, the growth of a country through income has both a negative and positive impact on the labour market. However, at a micro-level, the participation of female is determined by social norms. The chapter also includes the conceptual framework that gives a precise discussion of the relationship between FLFP and economic growth. The conceptual analysis postulates that the relationship inverts a U-shape. The next

chapter focuses on South African and international empirical literature on the determinants on the labour force participation rates.

## **CHAPTER 3: SOUTH AFRICAN LABOUR MARKET AND EMPIRICAL LITERATURE**

### **3.8. INTRODUCTION**

Various theories have been discussed to present the FLFP function which can be estimated through the use of econometric models in Chapter 2. Economists, Williams and Becker (1964, 1975) and Sloane et al. (2013), discussed the nature of the relationship that exists between labour force participation and its determinants. Moreover, the direction to which each determinant negatively or positively affects the participation in the labour force has been investigated both nationally and internationally. This chapter reviews both South African and international empirical literature on the determinants on the labour force participation rates in order to develop appropriate econometric models to be assessed in the next chapter.

The first section of the chapter discusses the South African labour force followed by the empirical literature. The main intention in discussing the South African labour market is to have an idea of what has been happening in the labour market over the last few decades. The chapter is designed as follows: section 3.2 discusses the South African labour market while section 3.3 discusses the empirical literature which is divided into two sections: South African literature in section 3.3.1 and literature from abroad in section 3.3.2 respectively.

### **3.2 THE SOUTH AFRICAN LABOUR MARKET**

Generally, labour market performance in South Africa has been relatively poor but developed after the promulgation of labour legislation during the post-apartheid era: The Basic Conditions of Employment Act of 1997 which advances the development of the economy and social justice through providing for fair working conditions; the Labour Relations Act No.66 of 1995 which also promotes development, peace and democracy in the workplace; and the Employment Equity Act No.55 of 1995 that protects workers and job seekers from unfair discrimination in the labour market.

Despite this legislation and regulations implemented to reduce the number of discouraged workers and promote disadvantaged groups to enter the labour force since the advent of democracy, the labour market has been characterised by an

increasing unemployment rate. The low labour market performance has been due to various issues which include militant trade unions, inflexible labour policies, high wages cause by the central bargaining system, and low and poor levels of education. With regard to education, Martin et al. (2012) postulate that South Africa is the worst performing country in terms of mathematics and science compared to its peers globally. However, the South African education system continues to produce graduated job seekers, which leads to increasing labour force participation.

Festus et al. (2016) state that the introduction of the labour legislation post 1994 limits employers from adjusting their labour demand while promoting workers' rights. Moreover, although the labour market has become more efficient, the rate of unemployment has increased. The economy is facing a serious problem in the labour market, in that the formal sector and informal sector are not equal (Festus et al., 2016). The rising rate of unemployment and participation restrict growth in the informal sector (Banarjee et al. 2008).

The South African economic growth has been unable to accommodate or create job opportunities for the labour force participants and has a high population growth rate which contributes to the rapidly rising unemployment rate. A study conducted a few years after democracy was instituted (Bhorat & Oosthuizen, 2006) indicates that the employment created after democracy cannot accommodate all the labour force participants. Employment creation and economic growth could not reduce the rising rate of unemployment (Festus et al., 2016). Nevertheless, Yu (2008) stipulates that the low skilled and uneducated labour are more likely to be employed in the agricultural sector within the country, for that reason, the participation rate increases together with the growth of the economy through the agricultural sector.

### **3.2.1 The Labour Force: The Level of Unemployment and Employment in RSA**

South Africa currently has a 27.5% unemployment rate (South African Reserve Bank [SARB] 2018) which is very high and the biggest challenge the country is facing. There has been a decrease in the level of economic growth and other economic fundamentals, so the unemployment rate has increased over several decades. In terms of the narrow definition of unemployment, (i.e. the people who are unemployed but are seeking jobs and trying to get some work) recent figures indicate a rise in the

jobless rate from 9.24% in 1980 to 27.7% in 2018 (SARB, 2018) meaning that one out of four South Africans in the labour force has failed to find work. The broad definition (i.e. both job seekers and people who are unemployed and available to work but have not taken active steps to look for work) indicates that it has increased from 29% to 40% (Hodge, 2009). The unemployment problem results directly from labour market issues such as wage rigidity and structural shifts in the economy (Festus et al., 2016). The economy consists of eight main sectors: manufacturing, retail, financial services, communications, mining, agriculture, tourism and the informal sector. The most noticeable change in employment over the last two decades has been the increase in the jobs in the financial sector and a decline in the agricultural and manufacturing sectors. These shifts show a move from a labour-intensive to a capital and skills-intensive economy.

The poorly performing educational system has affected the labour market negatively and has worsened the challenge of finding jobs faced by work seekers, Moses (2016). The static level of employment and increasing rate of unemployment induce small firms to pay lower wages and, at the same time, they take on less experienced workers which reduces labour market opportunities for more highly qualified people who are willing to work and are searching for work. Levels of self-employment in South Africa are also extremely low by international standards. Although labour productivity is rising in South Africa, this is driven by a process in which employers are shifting to higher productivity production techniques (relying on more skills and equipment) and through the letting go of firms' least productive employees, Rankin (2016).

### **3.3 EMPIRICAL LITERATURE**

At an empirical level, various theories have been used by different researchers to specify FLFP functions which were estimated using econometric techniques. Theories which indicate the nature of the relationship existing between FLFP and its selected determinants have become an empirical issue in most developing and developed economies. To justify the main determinants of FLFP and prior expectations, the empirical literature in South Africa and other countries is discussed.

FLFP has been observed as derived from the growth of a country. This section presents empirical evidence focusing on what has been done by other researchers in South Africa and abroad. Many studies conducted in various countries yield different

results. Different econometric modeling has been adopted to investigate and analysing the determinants and factors affecting FLFP of a country. These studies have been conducted based on the different choice of variables, sample size, and data frequency. There are survey-based, cross-sectional and time-series studies in the literature on the participation of women in the labour force. A number of studies have been conducted in developed and developing countries. Part 3.3.1 of this chapter addresses the literature which has investigated the phenomenon in South Africa. Section 3.3.2 focuses on studies conducted internationally using a mix of factors and methods to estimate the FLFP model.

### **3.3.1 South African Literature**

In the South African context, a few studies have been conducted on microeconomic factors that affect the participation of women in the labour market by Casale and Posale (2002), Ntuli (2003; 2007), Van der Westhuizen (2007) and Yakubu (2010). Different sample size, methodology, and variables were adopted in these studies. Other South African studies (Bharat & Oosthuizen, 2004; Oosthuizen, 2006; Yu, 2013) have been conducted to analyse the country's labour market and the labour force participation rate (LFPR) during the post-apartheid period.

Ntuli (2007) examined the determinants of the South African woman's labour force participation during the first decade of democracy which is the period of 1995–2004, by using survey data. The study adopted a decomposition technique used by Even and Macpherson (1990) which treats women's participation as a function of education, non-labour income, marriage, fertility and geographical variation in the development of the economy. The results state that the participation of female in the labour market is positively related to education while non-labour income, marriage, fertility and geographical variation in the development of economy lower their participation. A paper by Van der Westhuizen et al, (2007) reported the status of woman in the South African labour market for the period of 1995–2005, which found that feminisation of the labour market has been derived from large studies of African women's involvement in economic activities.

Yakubu (2010) investigated factors influencing FLFP in South Africa in 2008. He estimated a logistic model by using data collected from the 2008 Quarterly Labour

Force Survey conducted by Statistics South Africa. The study focuses on HCT which emphasises that women's education is positively related to their participation in the labour market, as education attainment increases the number of women in the labour force. The findings state that there is a positive association between FLFP and the level of education; hence, the findings were theoretically supported.

Casale and Posel (2002) investigated the relationship between female labour force and economic growth. The researchers estimated the interrelationship of these two variables (FLFP and economic growth), but they were not able to come up with an appropriate answer, due to unavailability of data when trying to measure FLFP in South Africa. During the 1990s, South African researchers (Bharat and Oosthuizen: 2004; Ntuli and Wittenberg: 2013) tried to look at this relationship by focusing on several different variables such as an increase in education, decrease in fertility rates and decrease in the number of marriages without including other variables like unemployment and female population as was done in other countries, which was the reason why it was not easy to trace the relationship.

A study conducted by Ntuli and Wittenberg (2013) postulates that the increasing level of education of black female in South Africa and urban residence were associated with a higher rate of women participating in the labour market. However, non-labour income, marriage, and fertility are associated with low rate of black women's involvement in the labour force. The study investigated changes that occurred in black South African women's labour force participation by employing a regression analysis of cross-sectional data for 1995, 1999 and 2004 respectively.

In South Africa, the following studies were conducted to analyse the country's labour market and the LFPR during the post-apartheid period. Bharat and Oosthuizen (2004) investigated changes that occurred within South Africa that brought about changes in the labour market in the post-apartheid period between 1995 and 2004. Later on, Oosthuizen (2006) updated Bharat and Oosthuizen's work by using the Stats SA labour force survey (LFS) of September 2004 and the Stats SA October Household Survey (OHS) of 1995. The results indicated that total employment increased while the unemployment rate rose over time, which indicates that the economy demonstrated the phenomenon of jobless growth – rising GDP without a concomitant rise in the employment rate. Moreover, the authors found that unemployment was

growing among the educated labour force which may be attributed to a skills mismatch and scarce skills shortages faced by the country. Oosthuizen's (2006) findings are in line with the work of Bharat and Oosthuizen (2004). The probability of an individual finding a job depends on level of education and gender. Gender appears to have an impact on the participation in the labour market.

### **3.3.2 Empirical Evidence from Abroad**

Numerous studies with regard to women's participation in the labour force have been conducted even at an international level. There is an abundance of empirical evidence based on female participation in the labour force using models applied within various country-specific contexts. Similar to national studies, international studies have been based on the different choice of variables, sample sizes, and methodological approaches. There are also longitudinal, cross-sectional, panel and time-series studies (Sackey, 2005; Tsani et al., 2012; Thévenon, 2013; Yakubu, 2010).

#### **3.3.2.1 Variable mix in different countries by different researchers**

FLFP depends on macro and micro economic factors and on different sectors of the economy. Researchers have used different variables to estimate FLFP in different countries such as Pakistan (Fatima & Sultana, 2009). At a macro level, Lechman and Kaur (2015), Elmi (2013), and Lechman (2014) investigated the participation of women in the labour market considering a country's economic growth, unemployment, and educational attainment.

Lincove (2008), working with data from various developed economies, found a link between economic growth and female participation in the labour market participation. Mujahid and uz Zafar (2012) investigated the relationship between FLFP and economic growth in Pakistan over a period of 30 years. Forgha and Mbella (2016) investigated the determinants of FLFP in Cameroon over a period of 37 years, adopting the generalised method of movement technique. Both studies used economic growth as their explanatory variable. Both Lincove and Mujahid and Zafar found that there was a long-run relationship between FLFP and economic growth while Forgha and Mbella indicated an opposite significant effect between female participation and economic growth.

Gaddis and Klasen (2014) re-estimated the FLFP which follows a U-shape trend, indicating that, as countries develop due to structural changes, more women participate in economic activities. FLFP is treated as the dependent variable and explanatory variable is GDP per capita and its square, education. In their findings, they state that there is no substantial empirical evidence of a systematic U-shape relationship between FLFP and GDP, while they also discovered that structural shifts were correlated with female participation in the labour force.

Tam (2010) investigated the U-shape female labour participation with economic development using panel data from 130 countries from 1950 to 1980. The study performed a regression analysis using female labour participation rate as the dependent variable and LGDP and LGDP<sup>2</sup> as the explanatory variables, and the OLS estimate was applied on the pooled data. The results suggest that the U-shape pattern shows up as an intertemporal relationship, thus this relationship cannot be taken as evidence in a time-series analysis.

Lechman and Kaur (2015) and Lechman (2014) provided new information about FLFP and economic development by investigating the relationship that exists between the two economic variables. A cross-country panel data of 162 countries over a period 1990-2012 was used. Assuming non-linearity between variables, the analysis considered only two different variables, the dependent variable, FLFP, which explains the share of women between 15 and 64 years of age in a nation's total labour force. The level of economic development was an independent variable which is the GDP per capita and the national log of purchasing-power-parity adjusted the national GDP per capita. All the data were taken from the World Bank's World Development Indicators of 2013. The results confirmed the existence of the U-hypothesis.

Tsani et al. (2012) investigated the relationship between FLFP and economic growth in southern Mediterranean countries using the general equilibrium model (GEM-E3-MEDPRO model). The cross-section data for 160 countries was used for the period 1960- 2008. FLFP was the dependent variable which was defined as the number of female workers aged 15 to 64 divided by the total female population of the same age group; independent variables used were real GDP and its square, and a set of controlling variables, which included education, fertility, urbanisation and the unemployment rate and dummy variables as their independent variables. The

researchers discovered that social grants and the number of children the women had had a negative effect on FLFP while education had a positive impact. Other controlling variables such as urban population were found to have no influence on FLFP; hence, they concluded that female participation in the labour markets of the Mediterranean countries was lower than other developed and developing countries

Fatima and Sultana (2009) using panel estimation approaches, explored the existence of the U-shape relationship between female participation in the labour force and economic development in the case of Pakistan using a simple fixed effects test for provinces and regions pooled for three years. The study utilises FLFP as their dependent variable. The independent variables were a vector of variables representing personal characteristics e.g. education and wages, a vector of variables for labour market conditions and the level of economic development in the country. The findings confirmed the U-shaped relationship, namely that higher economic development induces women to enter the labour force beyond a threshold level.

Chamlou, Muzi and Ahmed (2011) investigated the determinants of female labour supply in the Middle East and North Africa (MENA) countries. The study focused on the role of social norms and on the level of higher education as the drivers of women to enter the workforce. The findings were theoretically supported by the HCT indicating that education was positive and statistically significant while social norms indicated a negative effect on FLFP.

Furthermore, Elmi (2013) discussed the effects of economic development on FLFP rates and analysed its determinants in selected countries within the MENA group for a period of 20 years from 1990 to 2010 within the framework of U-shaped hypothesis, by using the method of Fractional Panel Probit. In this model, FLFP rate was treated as the dependent variable, while the level of higher education, unemployment and fertility rate were the independent variables. The estimation results suggest that relationship between economic development and the rate of female participation in MENA was U-shaped. The results were the same as Chamlou's findings, namely that higher education has a significant and positive effect on the female participation rate. Unemployment and the fertility rate reduce female participation rate in labour force.

At a micro level, Bhalla and Kaur (2011), Patel (2012), Ejaz (2007) used household size, marital status and family earnings as drivers of FLFP. These studies used survey

data to examine and analyse the determinants of female labour supply in emerging countries. Mafune (2013) examined factors that influenced women to participate in the labour force in Namibia. The data was collected from the Namibian Household Income and Expenditure Survey of 2009/2010. The participation of women was evaluated in terms of the level of education, age, and marital status. The results revealed that education was positively significant while marital status had a negative correlation. The findings were in line with the study conducted by Warunsiri and McNown (2010) which investigated the supply of female labour in Thailand for a period of 1985 to 2004 using the data collected from the Thai LFS. These two studies used similar variables (education, age, and marital status) to study the participation of women in the labour market.

Bhalla and Kaur (2011) examined the determinants of FLFP in India. The researchers used panel data analysis over a period of 17 years from 1993 to 2010. The FLFP rate was regressed as a dependent variable against explanatory variables which included education, female to male ratio, population and average household size. In a preliminary study, Contreras, Puentes and Bravo (2005) adopted a synthetic cohort method to analyse the evolution of female participation in Chile. The explanatory variables were education, the number of children in different age categories and the number of family members who were working. These variables affected the decision to participate throughout all stages of the lifecycle of women. The findings indicated that educational variables were positively correlated with participation. The macroeconomic settings may influence female participation. Therefore, such variables should be considered as factors determining the decision to work and their impact on GDP growth.

Ejaz (2007) identified the major determinants of FLFP in Pakistan, specifically with reference to rural and urban areas. Limited dependent variable techniques (Logit and Probit) has been applied to determine the factors affecting FLFP. This analysis used the data taken from the PSLM (Pakistan Social and Living Standards Measurement Survey, 2004-05) which measure individual and household characteristics of women between the ages of 15 and 49. FLFP was the dependent variable whereas the explanatory variables included age, education, marital status, the size of the family, household income and dummy variables. Empirical results suggested that age,

educational attainment, and marital status have significant and positive effects on FLFP.

### **3.3.2.2 Methodological approach**

The participation of women in the labour force has been modeled and estimated in various ways with different exogenous variables. However, it is notable that some studies are based on survey data, time-series, cross-sectional and panel data. Different methodologies, data frequency, sample size and econometrics techniques have been adopted. The methodology adopted by each study is determined by the type of the study and the availability of the data.

In emerging countries, multinomial logistic models have been applied in different studies by different researchers. Ateino (2006) in Kenya and Sackey (2005) in Ghana investigated the factors that determine or lead to female participation in the labour market. The studies were conducted based on the agricultural sector, but the results were not comparable because of different climate conditions in the countries. Sackey indicated that more women participated in the labour force in Ghana because the agricultural sector is dominant in the economy. Ateino revealed that education increased women's chances to participate in the private sector. Moreover, an individual's location and household headship had a significant effect on the participation of women.

Other studies in developing countries also employed probit and logit models to investigate the relationship between FLFP and economic growth using different types of data. Pieters and Klasen (2011) in India and Chen, Shao, Murtaza and Zhao (2014) in China examined the drivers of FLFP in urban and rural areas. Klasen and Pieters found that cultural and social factors influenced female participation rate more strongly than other factors while Chen et al. indicated that education was the main determinant in both rural and urban areas for China. Siphambe and Motswapong (2010) observed the major determinant of female participation in the labour market using the same method as the above-mentioned researchers. The results which found that education drives women to participate in the labour market in various areas of Botswana supported Chen et al.'s findings.

Mauritius is one of the developing countries where Tandrayen-Ragoobur, Ummersingh and Bundhoo (2011) studied the factors influencing women's participation in the labour force via the estimation of a logistic regression model. Primary data was collected via the household survey for the period from 2006 to 2008 and was used to model FLFP against exogenous variables which included the level of education, population growth rate and marital status. Their findings revealed that the higher the level of education a woman had, the higher the level of participation, while marital status had a significant effect because married women were found to be less likely to be part of the active population in the labour force. The results are in line with the study conducted by Ateino (2006) in the Kenyan economy.

The studies by Mazalliu and Zogjani (2015) and Sorsa, Mares, Didier, Guimaraes, Rabate, Tang and Tuske (2015) indicated a negative significant correlation between economic growth, the level of education and income share of women participating in the labour force. Mazalliu and Zogjani studied the effect of women's engagement in the labour force in various economic sectors, barriers and their contribution to a country's growth by utilising secondary data for Southern Eastern European countries. Sorsa et al. on the other hand, studied the determinants of female's participation in the labour market in India in comparison with OECD countries. Their results were similar to those of Mazalliu and Zogjani (2015).

Abdullah and Bakar (2011) investigated the causal relationship between total fertility rate and FLPR in four ASEAN countries for the period of 1980-2008, respectively. The study adopted the Engle-Granger test for causality. Specifically, a unit-root test was performed for all variables using an ADF test. To confirm the results, cointegration tests were employed to determine whether there was a long-run relationship among the variables, using the Johansen test. Their findings indicate that there is a long-run relationship among variables in the four countries.

Thévenon (2013) analysed the drivers of female participation in the labour market using a country's data for a period from the early 1980s for 18 OECD countries. The findings were in line with the study conducted by Mafune (2013) and Warunsiri and McNown (2010) which indicated that the increase in female education had increased the level of their participation in the labour force. Moreover, Sudarshan and Bhattacharya (2000) conducted a study based on women's participation in the labour

force in Delhi between September and November 2006, using the same variable mix. The results were similar to those of Thévenon's work.

Evans and Kelley (2007) Investigated trends in women's labour force participation in Australia for a period of 18 years from 1984 to 2002, using pooled cross-sectional data collected from the International Social Science Survey / Australia. LFP was employed as the dependent variable and education, marital status and the years of the survey as independent variables. The findings indicated that education had a positive effect while marital status had a negative effect; this is similar to the numerous studies mentioned above.

The low rate of women's participation in the labour force in India motivated Das, Jain-Chandra, Kochhar and Kumar (2015) to examine what determined the participation of female in the labour force in India. The study modeled the participation of women in the labour force based on demographic conditions and education. They found that most policies motivated or supported female economic participation (labour market flexibility, investment in infrastructure).

Amador, Bernal and Peña (2013) investigated the rise in female participation in Colombia over a period of 22 years from 1984 to 2006. This study was conducted using a cross-sectional data collected from the Colombian Household Survey, analysing women between the ages of 18 to 65 in the 10 largest cities. Female participation was investigated by the explanatory variables, which included education, marital status, and fertility rate. The findings were that education level, marital status, and fertility influenced the rate of female participation.

Blau and Kahn (2007) investigated married women's labour supply behaviour over a period of 20 years from 1980 to 2000 using cross-sectional data collected from the CPS. The labour supply was studied using an OLS and Tobit method. The independent variables were family income, family assets, spouse's hourly wage and a vector of controlled variables while the dependent variable was the hours worked. The findings revealed that married woman's real wages increased in both the 1980s and 1990s which caused an increase in labour supply of women.

In the construction industry, Hudu (2014) investigate the participation of female in the construction industry. The study utilised lack of interest, poor image and sexual

harassment as explanatory variables which are also found to be the reason why women are dominated by males in the construction industry. Using the Structural Vector Autoregressive (VAR) model, Kinoshita and Guo (2015) analysed the factors that could boost FLFP in Asia. The findings indicate that cash allowances reduced female employment whereas the gender wage gap encouraged employment.

### 3.4 SUMMARY OF EMPIRICAL EVIDENCE

Table 3.1 below indicates the variables used to determine the relationship between FLFP, the labour market and its determinants, and the techniques and models used in various studies.

**Table 3.1: Selected evidence from South Africa**

Reference	Dependent variable	Explanatory variables	Models and techniques
Ntuli & Wittenberg (2013)	Women labour force	Marital status, level of education	Survey logit estimate
Yakubu (2010)	Female labour force	unemployment	Logistic regression model
Festus et al. (2016)	Labour force 1995-2015	Unemployment & employment	OLS
Floro & Komatsu (2011)	Labour force	Family characteristics, fertility	Probit analysis
Casale & Posale (2002)	Female labour supply	Education	Survey logit estimation
Ntuli (2007)	Labour force participation 1995-2004	Age, marital status, province	Logit modeling & decomposition technique
Oosthuizen (2006)	Labour force	Province, education	Logit and probit model
Van der Westhuizen (2007)	Labour force 1995-2005	Race, province	Probit model

The empirical literature indicates that women during apartheid were deprived of participating in the labour market but were generally housewives and dependent on their husband's income, working only on family farms. The post-apartheid government introduced policies that encourage women to participate in the labour market and to become better educated. Affirmative action policies also increased women's employment and discouraged discrimination in the labour market. Incentives like

Isivande Women’s Fund and the Women’s Entrepreneurial Fund offered to women accelerated women’s participation in the labour market.

**Table 3.2: Evidence from the rest of the world**

Reference	Country	Dependent Variable	Independent Variables	Model Specification
Abdullar and Baker (2011)	ASEAN countries 1980-2008	FLFP	Total fertility	ADF test
Amador, et al. (2011) Colombia	Colombia	FLFP	+Education, Marital status Fertility	Logit and probit model
Blau and Kahn (2007)	Different countries	Labour supply	Spouse hourly income, Family income and assets	OLS & Tobit model
Chaudhuri (2010)	Different countries	FLFP	Real GDP per capita GDP	Multivariate regression analysis
Chen et al. (2013) 2003-2010	China	Family Utility	Consumption	Probit model
Contreras, et al. (2005)	Chile 1987-1997 annual data	FLFP	Age rate, Population Consumption	Synthetic cohort analysis
Ejaz (2007)	Pakistan 2004-2005	FLFP	+education, +marital status,+age	Logit and Probit
Elmi and Shalmani (2013)		FLFP	Education	OLS model
Fatima (2009) Parkistan	Pakistan	FLFP	+education, unemployment, +wage rate	OLS model
Forga & Mbella (2015)	Cameroon	FLFP	Per capita income, fertility rate, male labour force	Generalized method of movement technics
Gaddis and Klasen (2013)	Different countries	FLFP	Structural changes, GDP per capita (PPP), EDUCATION	OLS
Goksel (2013)	Turkey	FLFP	Education, fertility	OLS regression

Reference	Country	Dependent Variable	Independent Variables	Model Specification
Hudu (2014)	Asia	FLFP	Sexual harassment, lack interest & poor image	Structural VAR model
Ince (2010)	Turkey	FLP	GDP, unemployment, fertility & literacy rate	Time-series regression
Jaumotte (2013)	OECD countries	Female Participation	Education ,, Marital status ,, Taxes	Logistic regression
Kinoshita and Guo (2015)	OECD countries	Female labour supply	Child care, child benefit	Structural VAR
Lahoti & Swaminathan (2013)	China 1983-2010	FLFP	GDP, wage rate, education	OLS
Lechman (2014)	162 countries	FLFP	Economic growth	Longitudinal analysis, IV regression
Lechman and Kaur (2015)	162 countries	FLFP	Real GDP	OLS ordinary least squares (OLS)
Mafune (2013) Namibia	Namibia	FLFP	+education, -marital status, +Age	Probit regression
Mujahid and Zafar (2012)	1980-2010 annual data Cameroon	FLFP	Economic Growth, real GDP per capita	OLS
Patel (2012)	India	FLFP	Population, sex ratio, female education	OLS analysis
Roopnarine & Ramrattan (2011)	Pakistan	FLFP	Level of schooling, age, non-labour income, marital status	Probit model
Sarsa, et al. (2015)	India	FLFP	Individual, household & state characteristics	Probit & OLS method
Siphambe & Motswapong (2010)	2005-2006	FLFP	Marriage status, location (urban/rural)	Probit & Logit model
Taamouti & Ziroili (2011)	Morocco	FLFP	Age, age squared, marital status	Logistic regression
Tam (2010)	130 countries	FLFP	LGDP, LGDP <sup>2</sup>	OLS regression

Reference	Country	Dependent Variable	Independent Variables	Model Specification
Tandrayen-Ragoobur, et al (2011) Mauritius	Mauritius	FLFP	+education, -marital status	Logistic regression model
Théevenon (2013)	OECD countries	FLFP	GDP, unemployment, child care, part-time work	Pooled OLS
Tsani, et et al. (2012)	Various countries	FLFP	GDP, real GDP per capita, education, fertility, urbanisation, unemployment	General equilibrium model
Warunsuri and McNown (2010)	Thailand 1985-2004	FLFP	Employment, hours of work	Weighted least squares
Wusu (2012)	Nigeria	Fertility rate	Female education, FLFP	OLS regression

Studies from the rest of the world confirm the U-shape hypothesis (Forga & Mbella, 2011; Lechman & Kaur, 2015; Tam, 2010; Tsani, 2012) as there is overwhelming evidence of increasing FLFP based on the economic structural shifts. During the early stages of economic growth where there was limited access to education, the majority of workers were low skilled and participated in the agricultural sector. Forga and Mbella (2016) postulate that this implies that at the initial stage of growth, female labour participation is relatively high but as the economy proceeds along traditional growth patterns where the industrial sector is gradually substituted for the agricultural sector, female participation falls.

### 3.5 CHAPTER SUMMARY

Empirical studies indicate that the impact of education on FLFP in both developed and developing countries is theoretically supported (Amandor et al, 2013; Evans & Kelley, 2008; Ntuli, 2007; Yakubu, 2010). Investment in human capital through education increases the participation of women in the labour force. The U-shape hypothesis was confirmed (Lechman & Kaur 2015; Tam, 2010; Tsani et al., 2014). At a low level of economic growth, low husband income and when a household breadwinner becomes unemployed, women supply more labour in the labour market to maintain their standard of living (Forgha & Mbella, 2016).

## **CHAPTER 4: STATISTICAL ESTIMATION METHODOLOGY**

### **4.1. INTRODUCTION**

This chapter discusses the theoretical and conceptual issues surrounding methodology of time-series data analysis, data issues and empirical regression modeling. Moreover, it presents how the study analysed the data to address the research objectives. The data issues section discusses the data sources, choice of variables and justification for the sample size. This chapter also shows how the time-series data is analysed through the use of single equations, Fully-Modified Ordinary Least Squares (FM-OLS), DOLS and CCR, as well as through the application of multivariate equation techniques which include the Vector Autoregression (VAR) and VECM (VAR-VECM) frameworks followed by the Granger Causality and impulse response function.

The chapter is organised as follows: section 4.2 explains the data source, justification of sample size, frequency and period of choice of variables as selected to estimate the model. Section 4.3 covers the data transformation and descriptive analyses of the various time-series observations. The spurious regression concept and unit-root tests are articulated extensively in section 4.4 and 4.5 respectively. Section 4.6 discusses the why the specific lag was chosen to perform a cointegration analysis. The cointegration analysis is presented in detail in section 4.7 which is followed by the FM-OLS and the VAR-VECM analysis in section 4.9. Lastly, in section 4.10 the diagnostic inspection is presented.

### **4.2. DATA ISSUES**

This section gives full discussion based on extraction of data, description of variables incorporated in the model, and justification for the variable mix. The first section concentrates on extraction of data, outlines the process followed in getting the data, and gives a justification for the sample size. A full description of the variables included in the model is given in section 4.2.2 which is followed by the last section (4.2.3) that highlights the motivation for the variable mix.

#### 4.2.1. Data Sources and Sample Size

In order to meet its objectives, the study utilised two different sets of time-series data: the yearly frequency from 1980 to 2015, and the quarterly frequency from 2008 to 2016.

All the quarterly data and most of the yearly data, except for the annual FLFP Rate were downloaded from online statistical query in the South African Reserve Bank website (<http://www.resbank.co.za>) using relevant KBP codes. This data is also available on the World Bank website. The South African FLFP yearly data was extracted from the United Nations conference on Trade and Development (UNCTD) database, although the sample size is relatively small it is large enough not to expose the study to the micro-numerosity problem. Micro-numerosity is seen to be one of the causes for regression pathologies such as multicollinearity; thus, the researcher should consider this problem to maintain the accurateness of the model. The data in all variables are expressed in 2010 constant prices in South African Rands, except for FLFP which is expressed as a percentage.

#### 4.2.2. Description of Variables

The study mainly seeks to investigate the long-run impact of economic growth and unemployment on FLFP. The model specification incorporates variables as follows: FLFP expectations as an indication of female labour force employment and participation in the economy of the working population of working age (15-64 years) that are economically active; GDP per capita at constant prices as a proxy of output which reflects the level of productive of the economy; and unemployment as the macroeconomic indicator affecting the economy as a whole. All variables will be abbreviated in the model, FLFP output (GDP), unemployment (UNEMP).

**Table 4.1: Summary of description of variables**

Variable	Description or definition	Metric	Source
FLFP	Female labour force participation expectation	Percentage	UNCTAD (yearly) STATSSA (quarterly)
GDP	Output, real GDP at 2010 constant prices net of export	Rands per capita	SARB KBP 6006y

Variable	Description or definition	Metric	Source
UNEMP	The narrow definition of the unemployment rate (actively seeking employment within the last three weeks)	Percentage	World Bank (yearly) SARB KBP7019K (quarterly)

#### 4.2.3. Justification of Variable Mix and Sample Size

The present study employed variables that were adopted by Lechman and Kaur (2015) in verifying the feminisation hypothesis for 162 countries. This study modified the Lechman and Kaur (2015) specifications slightly, to include the unemployment rate as a possible driver of female participation rate in the economy where high unemployment arising through high reservation wages required by male workers incentivises employers to employ more women at a lower wage rate. All variables utilised in this study have been used frequently nationwide in empirical literature by various researchers (Lechman, 2014; Tam, 2010; Tsani et al., 2012). For instance, South African studies (Ntuli, 2007, 2013; Yakubu, 2010) focused only on survey data which did not include the overall output of a country. Lechman and Kaur (2015) utilise similar economic variables but excluded the effect of unemployment.

Time-series data set of 36 yearly observations is used from the period from 1980 to 2015 and 33 quarterly observations from 2008 to 2016. Although the sample size is relatively small, it is scientifically acceptable (Mutsau, 2012). The small sample size has been influenced by the unavailability of the data of FLFP and unemployment rate. However, as mentioned above, the study utilises the expected FLFP rate as a proxy for female participation in the labour force since there is no data available from the past two decades specifically for FLFP.

#### 4.2.4. Data Transformation and Descriptive Analysis

Data transformation of raw observations is necessary to make the variables amenable to econometric modeling. In this study, transformations included conversion of all variables that are in nominal terms to real terms, to avoid artificial trends that might arise due to nominal changes induced by inflation. Furthermore, real GDP was first transformed into natural logarithms and an added variable was created by squaring it so that the existence of the U-shape hypothesis in FLFP rate relative to GDP could be tested. Variables in levels were converted to first differenced (FD) variables to enable

the researcher to apply the Granger Causality test. With regard to the regression analyses, the main intention of running logarithmic transformation is to enable the researcher to express coefficients in elasticity values and mitigate regression problems such as serial correlation and heteroscedasticity.

After all transformations, visual inspection of data by plotting line graphs and descriptive statistics as part of preliminary analysis was done. Graphical representation of variables generally enables the researcher to have closer inspection of possibilities of structural breaks that might exist among variables in consecutive years, and it becomes easier to identify outliers. According to Gujarati and Porter (2009), graphs play a significant role in providing clarity of an anticipated nature of the series with respect to inclusion of trends or constants, or both in the model. Descriptive statistics include measures of central tendency and dispersion (mean, mode and median) which would assist in understanding the distribution of each variable in the data set. For this objective, relevant calculations of the mean, mode, and median were computed and portrayed in table form. Furthermore, measures of skewness and kurtosis which would enable the researcher to understand deviations of variables from normal distribution and any other outliers that might exist in the data set were computed by means of the Jarque-Bera (JB) test of normality as well as box-and-whisker plots. According to Mutsau (2012), the kurtosis has a coefficient of 3 while skewness has zero as its coefficient; both these coefficients usually apply in normally distributed data sets. The correlation covariance matrix was applied to understand co-movements of variables and anticipated signs of coefficients. This, in turn, helped the researcher formulate prior expectations before running the regression.

#### **4.3. SPURIOUS REGRESSION CONCEPT**

The problem of spurious regression requires serious attention to ensure that the accuracy and reliability of findings of the study are not compromised. This time-series problem emanates from the non-stationarity of data. It is also argued that spurious regression could arise in non-stationary series despite a large sample. In the presence of spurious regression, statistical tests are likely to reflect higher than necessary values even when variables applied in the analysis have no clear interrelationships. This may indicate a positive relationship that does not exist between variables. Granger (1988) proposed a rule of thumb that can be utilised to detect spurious

regression. According to the rule of thumb, the OLS regression is said to be spurious if  $R^2$  exceeds the Durbin Watson (DW)<sup>1</sup> statistic or if  $R^2$  is approximately 1. This implies that an extremely low DW statistic reflects the presence of strong first-order autocorrelation.

### 4.3.1. Stationarity of Time-Series

The stationarity of time-series indicates that its statistical properties such as variance, mean, mode and autocorrelation are all constant (Asteriou & Hall, 2015; Gujarati & Porter, 2009; Mutsau, 2013). Time-series is assumed to be rendered stationary through transformation which makes it easy to predict. Mostly, economics and business series are found to be non-stationary in their levels. However, even after seasonal adjustment of high frequency variables, some series are likely to exhibit trends, non-stationarity or random walk behaviour. If the series indicate a long-run trend, then it may be stationarised by de-trending (Gujarati & Porter, 2009). Thus, in such a case, the series is said to be trend-stationary. However, in some cases it happens that de-trending does not stationarise the series; then, it is necessary to transform the series to season to season or period to period so that it would be difference stationary. Non-stationary series need to be stationary to obtain meaningful results or sample statistics such as mean or variance. These statistics are more useful in predicting the future behaviour of a variable if the series is stationary.

## 4.4. UNIT-ROOT TEST

Unit-roots refer to the existence of stochastic trends among series due to time variant means and variance of series which render them to be non-stationary. It follows that stationary time-series are only those with constant means and variances.

The constant mean is given by:

$$E(Y_t) = \mu$$

$$\text{Variance: } \text{VAR}(Y_t) = E(Y_t - \mu)^2 = \sigma^2$$

---

<sup>1</sup> The Durbin–Watson statistic is a test statistic used to detect the presence of autocorrelation at lag 1 in the residuals from a regression analysis

Covariance:  $\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$

where  $\gamma_k$  is auto covariance at lag (k). Violation of either requirements of stationary series would mean that variables are non-stationary; therefore, estimating regressions using traditional methods would lead to spurious regression. As indicated above, spurious regression tends to give high  $R^2$  and t-values than necessary. Thus, in maintaining reliability of findings of the study relevant transformation of variables from non-stationary to stationary is highly recommended. For that purpose, there are formal and informal tests of checking whether variables are stationary or non-stationary. Informal methods refer to visual inspection of data, correlograms and graphs. Although numerous tests could be applied, this study employed the Augmented Dickey-Fuller test and the Phillips-Perron test provided that informal tests are complementary to these formal tests.

#### 4.4.1. The Dickey-Fuller Test

This test was developed after Dickey and Fuller (1981) discovered that t-values of coefficient of lagged Y ( $Y_{t-1}$ ) under the null hypothesis that there is a unit-root follows a tau ( $\tau$ ) test. This originates from the fact that estimated coefficients of  $Y_{t-1}$  do not have an asymptotic distribution which renders the usual t-statistic invalid for the analysis and therefore creates the incentive to determine if the scientifically acceptable statistic could be used. Considering the following equation

$$Y_t = \rho Y_{t-1} + \mu_t \tag{4.1}$$

where  $\rho$  is a parameter ranges between -1 and 1, while  $\mu_t$  is a white-noise error term. This equation displays non-stationarity in the case of  $\rho = 1$  which is known as a random walk without drift, however, it requires further manipulation so as to enable us to regress it and apply relevant T-test. In its reduced form after subtracting  $Y_{t-1}$  from both sides of the equation it becomes:

$$\Delta Y = \delta Y_{t-1} + \mu_t \tag{4.2}$$

Where  $\delta = (\rho - 1)$  and  $\Delta$  is the first difference. The last equation should be in a form that accommodates that a random walk may have a no drift (Eq 4.3), have drift (Eq 4.4) and (Eq 4.4) have both deterministic and stochastic trends.

$$\Delta Y_t = \delta Y_{t-1} + \mu_t \quad (4.3)$$

$$\Delta Y_t = \beta_t + \delta Y_{t-1} + \mu_t \quad (4.4)$$

$$\Delta Y_t = \beta_1 + \beta_{2t} + \delta Y_{t-1} + \mu_t \quad (4.5)$$

To test for stationarity using this equation, the null hypothesis that there is a unit-root ( $H_0: \delta = 0$ ) and alternative which claims that opposites are set. Thus, rejecting the null hypothesis will mean that the variable in consideration is stationary. As indicated earlier, the estimated coefficients of  $Y_t$  do not have the usual t-distribution; therefore, critical values of MacKinnon (1996) are recommended in this regard.

#### 4.4.2. Augmented-Dickey-Fuller Test (ADF)

This test was developed by Dickey and Fuller (1981) as an extension of the Dickey-Fuller test discussed above. Basically, they extended the testing procedure to form an augmented version with the intention to remove autocorrelation by including extra lagged terms of the dependent variable. To select the lag length that constitutes the extra terms, one can use either AIC or SBC or any length capable to widen the residuals. Thereafter, residuals of the ADF regression can be tested to see whether they are autocorrelated or not through application of the LM test.

$$\Delta Y_t = \delta Y_{t-1} + \sum_{i=1}^p \beta_i \Delta Y_{t-1} + \mu_t \quad (4.6)$$

$$\Delta Y_t = \lambda_0 + \delta Y_{t-1} + \sum_{i=1}^p \beta_i \Delta Y_{t-1} + \mu_t \quad (4.7)$$

$$\Delta Y_t = \lambda_0 + \delta Y_{t-1} + \beta_{2t} + \sum_{i=1}^p \beta_i \Delta Y_{t-1} + \mu_t \quad (4.8)$$

It depends on the model to be included to see whether the first equation applies in testing a pure random walk. The other equations are related to testing a model of a

random walk with drift and a model with a deterministic trend respectively. The null hypothesis that a variable in consideration has a unit-root ought to be rejected or accepted based on MacKinnon's (1996) critical values.

In all these models, the ADF t-statistic is computed in respect of the following hypothesis.

$$H_0 = y_t \text{ has a unit - root}$$

$$H_a = y_t \text{ has no unit - root}$$

#### **4.4.3. The Phillips-Perron Test**

Phillips and Perron (1988) proposed a unit-root test which is an extension to the ADF test procedure that would allow fairly mild assumptions with regard to distribution of errors (Asteriou & Halls, 2015). In the same sense as the ADF test, the test regression for the Phillips-Perron (PP) test is first-order autoregressive (AR (1)) process,  $\Delta Y_{t-1} = \alpha_0 + \gamma Y_{t-1} + e_t$ . This test differs from the ADF test which corrects for higher order serial correlation through increasing lagged difference terms by correcting the t-statistic of the coefficient ( $\gamma$ ) from AR(1) regression. The main purpose of adjustments to the t-test is basically to address the problem of serial correlation in error terms ( $e_t$ ). Thus, the PP statistic is a modification of the ADF t-statistic which only considers the less prohibitive nature of the error process. The PP-tests are computed with three different models (i.e, intercept and trend or intercept only or with neither intercept nor trend) with the application of MacKinnon's (1995) critical values. The null and alternative hypotheses are the same as the ADF tests.

#### **4.5. LAG LENGTH SELECTION CRITERIA**

Brooks (2014) indicates that there are two approaches in selecting the true lag length in both multivariate and single regressions. Firstly, the lag information criteria is mostly applicable in economics studies in determining the optimal lag length that the research should use in testing cointegration. Secondly, the frequency of the data can be used in measuring and taking decisions on the selection of a correct lag length. However, this rule has limitations. In a case where the frequency is high (e.g., hourly), it may not

be applicable. If the time-series is daily, weekly or monthly, then the researcher is limited to 12 lags while on the quarterly and yearly frequency the lag should be 1 and 4 respectively (Mutsau, 2015)

Autoregressions (AR) have been indirectly and directly applicable in most economic studies. The application of such modeling needs the determination of the lag length selection criteria. The AR lag of time-series variables may be determined by many selection criteria that are applicable in most economic research. It is notable that the AR process of lag selection refers to a time-series which the current value depends on its previous  $p$  lagged values and it is denoted by AR ( $p$ ). AR ( $p$ ) has to be estimated using different lag length selection criteria since it is normally not known. Different selection criteria include (i) the Akaike Information Criteria (AIC) developed by Akaike (1973), (ii) Schwarz Information Criteria by Schwarz (1978), (iii) Hannan-Quinn Criteria (HQC) developed by Hannan and Quinn (1979), (iv) Final Prediction Error (FPE) by Akaike (1969) and lastly is the Bayesian Information Criteria (BIC) by Akaike (1979).

Liew (2004) indicates the important and disadvantage of each criteria. For large samples with 120 or more observations, the HQC is found to be superior and identifies the true lag length. However, this study utilises only 36 observations for single equations and 33 observations multivariate equations which is relatively small; therefore the study uses the AIC and FPE. These two criteria (AIC and FPE) are found to be more powerful and efficient than other criteria in selecting the correct lag length in a case whereby the sample size is small (60 and below) (Liew, 2004). Both AIC and FPE maximise the chances of recovering the correct lag length while minimizing the chances of underestimation of the lag length (Asghar & Abid, 2007).

#### **4.6. COINTEGRATION ANALYSIS**

Cointegration was first developed by Granger (1981) and, thereafter, Phillips (1967, 1987), Engle and Granger (1987) and Phillips and Ouliaris (1990) further elaborated on this topic. Engle and Granger (1987), in bivariate equations, elaborate more on the formal discussion of cointegration between two economic variables. Two variables ( $X$  &  $Y$ ) are said to move together  $\{I(1)\}$  where, for example,  $d \geq b > 0$ , notated as  $Y_t, X_t \sim C1(d, b)$ , if there is a linear correlation between variables and if both time-series are integrated.

Cointegration is an overriding requirement for any economic model using non-stationary time-series data (Asteriou & Hall, 2015). The regression becomes spurious and meaningless if variables do not cointegrate. Moreover, if there is a long-run association that exists among variables (X&Y), they are recognised as rising over time or they trend together in the same direction. Cointegration analysis helps to identify long-run economic relationships between two or more variables and to avoid the risk of spurious regression.

The AR is estimated using stationary data to avoid spurious results. If the regression is estimated using non-stationary data, this may indicate a relationship that does not exist (Gujarati & Porter, 2009). Cointegration as one of significant attributes of time-series variables usually applied to check for existence of long-run equilibrium relationships between two or more non-stationary variables. Two time-series variables (X and Y) are meant to be cointegrated if the order of integration is I(1), but residuals obtained from  $Y_t$  regressed on  $X_t$  are I(0) are stationary (Harris, 1995). Once cointegrating relationships among respective variables regressions have been identified, it is possible to analyse and interpret non-stationary variables. This follows the argument that a regression of cointegrated variables is not spurious

The concept of cointegration was developed intentionally to deal with the problem of estimating regressions using non-stationary which consequently render the model meaningless or spurious. One may decide to run the model with FD variables; however, such remedial to spurious regression situation is not recommended as it compromises long-run information, and changes the whole model to short-run (Harris, 2000). In order for two variables to have a long-run relationship, a linear combination of X & Y that is I (0) variable (Asteriou & Hall, 2015). The linear combination of both variables can be extracted from the following model;

$$Y_t = \beta_1 + \beta_2 X_t + \mu_t \quad (4.9)$$

$$\mu_t = Y_t - \hat{\beta}_1 - \hat{\beta}_2 X_t \quad (4.10)$$

Equation 10 is the estimation of residuals. If they are stationary, then variables are cointegrated.

Cointegration is typically presented as follows: when two variables (X & Y) that are I(1) and assume a vector that give a linear combination of both variables which is stationary:

$$\sigma_1 Y_t + \sigma_2 X_t = \mu_t \sim I(0) \quad (4.11)$$

The above equation indicates the cointegration set of variable X & Y and the cointegration vector  $(\sigma_1 \sigma_2)$ . Then the linear long-run relationship for  $Y_t$ ;

$$Y_t^* = \beta X_t \quad (4.12)$$

By normalising equation 4.11, we obtain the equilibrium or the long-run value of

$$Y_t; Y_t = -\frac{\sigma_2}{\sigma_1} X_t + \varepsilon_t.$$

Notice  $Y^* = -(\sigma_2/\sigma_1)X_t$  indicating the equilibrium (long-run) value of  $Y_t$ .

Cointegration of variables is tested in both bivariate and multivariate equations using different econometric approaches. There are two approaches to test for cointegration: single equation and multivariate techniques. The single equation comprises of Engle-Granger methodology introduced by Granger (1981) and Autoregressive Distributive Lag (ARDL) approach while the multivariate consist of Johansen methodology by Johansen (1988), and Bayesian econometrics. This study adopted the Johansen methodology.

#### **4.6.1. The Engle-Granger Approach**

For bivariate equations, the Engle-Granger approach is recommended to test for the equilibrium relationship among variables. The Engle-Granger approach to cointegration involves first ensuring that the variables are individually unit-root processes. Then a regression would be conducted of one of the series on the other and the residuals from that regression collected. These residuals would then be subjected to a DF or augmented ADF test. If the null hypothesis of a unit-root in the DF test regression residuals is not rejected, it would be concluded that a stationary combination of the non-stationary variables has not been found and that there is thus no cointegration. On the other hand, if the null hypothesis is rejected, it would be concluded that a stationary combination of the non-stationary variables has been found and thus that the variables are cointegrated. To understand this approach,

consider the following explanation of two series X & Y. If these variables are not trending in the same direction [I(0) & I(1)], then every linear combination will result in a series that will always be non-stationary. The non-stationary series will dominate the other (Asteriou & Hall, 2015). On the other hand, if both X & Y are non-stationary, the linear combination will be also be I(1). However Asteriou and Hall (2015) postulate that seldom happens, although some series indicate that the combination is I(0) and, in this case, we would say the variables are cointegrated of order (1,1).

The Engle-Granger (EG) cointegration test is easier to implement and understand but it has some shortcomings. First, in a case of more than two variables, the cointegration relationship may be more than one and the Engle-Granger approach is unable to detect such a relationship. It does not give a number of equilibrium levels when variables are more than two. Asteriou and Hall (2015) also stipulate that the EG test depends on a two-step estimation, which is to generate residuals and estimate their regressions to see if the series is I(0) or I(1). Moreover, estimating the residuals regression relies on the generated residuals. When estimating the long-run relationship using two (X & Y) variables, the test does not specify which variable should be endogenous or exogenous. For example, one may decide to use  $X_t$  as endogenous ( $X_t = \beta_1 + \beta_2 Y_t + \varepsilon_{t1}$ ) and  $Y_t$  as a dependent variable ( $Y_t = \beta_1 + \beta_2 X_t + \varepsilon_{t2}$ ). The EG test for cointegration indicates that the result of stationarity is the same for both regressions ( $\varepsilon_{t1}$  &  $\varepsilon_{t2}$ ). In economics, this is not true: the two regressions may exhibit different cointegration results ( $\varepsilon_{t1}$  may indicate that X and Y are cointegrated while  $\varepsilon_{t2}$  indicate the opposite). Furthermore, in cases of more than two variables, the problem become more difficult for the EG test to resolve. All the mentioned EG shortcomings are easily resolved by the use of multivariate cointegration test which is the Johansen approach discussed in the next section.

#### **4.7. JOHANSEN'S MAXIMUM LIKELIHOOD COINTEGRATION TEST**

At this point, it is important to clarify that the VAR methodology and the Johansen VECM approach were used to model the quarterly data exclusively, since the modeling of yearly data within the said frameworks did not produce robust results.

The second approach is the Johansen test which is best in determining the relationship of variables in multivariate equations. When testing for multivariate cointegration, one of the approaches has been to test for cointegration using a Vector Autoregressive

(VAR) approach. This assumes that all the variables in the model are endogenous, although it is possible to include exogenous variables as well, although these do not act as dependent variables. As with the bivariate cointegration case it is possible to produce long-run coefficients and error-correction models with this approach. It is called the Johansen Maximum Likelihood procedure. Not at all like the Engle-Granger approach, is that it is conceivable to have more than a single cointegrating vector. The test itself creates various insights which can be used to decide the quantity of cointegrating vectors exhibit. Another distinction with the Engle-Granger test is that there are two separate tests for the quantity of cointegrating connections and they do not generally agree with the number present. As indicated by Cheung and Lai (1993), the energy of the Johansen and Juselius' (1990) test is superior to that of the Engle-Granger test.

If the model has more than two variables, it is possible to have more than one cointegrating vector, Asteriou & Hall (2015). Moreover, variables might have different equilibrium relationships governing all variables in the model. For instance, n variables may result in up to n-1 cointegrating vectors. In a simple case where n equals 2, the cointegrating vector will be unique in the existence of cointegration. In a case whereby n is greater than 2 with an assumption that only one equilibrium relationship exists whereas there are more than one, then the problem cannot be solved by a bivariate cointegration test (Engle-Granger cointegration test). Therefore, in that case, an alternative approach such as the Johansen approach should be considered for multivariate regressions. The study has more than two variables that may indicate several equilibrium relationships governing the model. The study adopts the Johansen test of cointegration approach to resolve the cointegrating issue among the selected choice of economic variables.

To present this approach, the single equation is extended to multivariate equations. The model has three variables which can all be endogenous. This can be expressed using the matrix notation:

$$Z_t = A_1Z_{t-1} + A_2Z_{t-2} + \dots + A_nZ_{t-n} + \mu_t \quad (4.13)$$

Equation (4.13) is comparable to a single equation model in equation (4.12) for X and Y. However, it can be taken to a VECM approach as indicated below:

$$\Delta Z_t = \Gamma_1 \Delta Z_{t-1} + \Gamma_2 \Delta Z_{t-2} + \dots \dots \dots \Gamma_{n-1} \Delta Z_{t-n-1} + \Pi Z_{t-1} + \mu_t \quad (4.14)$$

Where  $\Pi$  equals  $-(I - A_1 - A_2 - \dots - A_n)$  and  $\Gamma$  equal to  $(I - A_1 - A_2 - \dots - A_n)$ . Eq (4.13) indicate a 3 by 3 matrix that should be examined and contains the equilibrium relationship between variables. For simplicity, this can be modeled as follows:

$$\begin{bmatrix} \Delta Y_t \\ \Delta X_t \\ \Delta W_t \end{bmatrix} = \Gamma_1 \begin{bmatrix} \Delta Y_{t-1} \\ \Delta X_{t-1} \\ \Delta W_{t-1} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \\ W_{t-1} \end{bmatrix} + \varepsilon_t \quad (4.15)$$

#### 4.8. THE VAR/VECM APPROACH

In economics, it is usual to have regressors that not only explain the given dependent variable but they can also be an endogenous variable and be explained by the variable which the regression claims to be dependent. The explanatory variable(s) can be also explained by the dependent variable at the same time (Asteriou & Hall, 2015). Therefore, we can say there is a bidirectional relationship between selected variables. According to Sims (1980), it is not easy to identify the variable that should be treated as the regressand or regressor in a case of simultaneous regressions. In the VAR model, each regression has the same set of variables as the single equation approach.

The application of a VAR modeling is treated symmetrically. For instance, a time-series may indicate that  $X_t$  is affecting  $Y_t$  while on the other hand  $X_t$  may be affected by the current and past series of  $Y_t$ . This can be expressed mathematically as follows;

$$Y_t = \beta_{10} - \beta_{11}X_t + \gamma_{11}Y_{t-1} + \gamma_{21}X_{t-1} + \mu_{yt} \quad (4.16)$$

$$X_t = \beta_{20} - \beta_{21}Y_t + \gamma_{21}X_{t-1} + \gamma_{22}Y_{t-1} + \mu_{xt} \quad (4.17)$$

The above regressions assume that both series are  $I(0)$  and residuals of both models are not correlated. Asteriou and Hall (2016) indicate that in such case the above equations are not in reduced form since  $Y_t$  and  $X_t$  have a contemporaneous impact on each other (given  $-\beta_2$  &  $\beta_1$ ). The matrix can be expressed as follows:

$$\begin{bmatrix} 1 & \beta_1 \\ \beta_2 & 1 \end{bmatrix} \begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \mu_{yt} \\ \mu_{xt} \end{bmatrix} \quad (4.18)$$

Or

In this study, for the sake of simplicity, the VAR is organised as follows:

$$Y_t = \delta + \Gamma_1 Y_{t-1} + \Gamma_2 Y_{t-2} + \varepsilon_t \quad (4.19)$$

The above equation is a second order VAR, where  $Y_t = (flfp_t, gdp_t, unemp_t)$  and is a  $(3 \times 1)$  column vector,  $\delta$  denotes a  $(3 \times 1)$  vector of constants and  $\Gamma_1$  a  $(3 \times 3)$  matrix of autoregressive coefficients. The VAR model cannot be estimated in its primitive form due to the contemporaneous correlation of error terms across equations hence the Cholesky decomposition was applied to the model resulting in the orthogonalisation of errors into a lower triangle Cholesky factorisation, which – with reference to the ordering of variables in the  $Y_t$  vector – implies that real output (the most endogenous variable) is assumed to be contemporaneously affected by all the other variables in the system, while budget deficits are contemporaneously affected by labour and capital and lags of output, and capital (the most exogenous variable) is affected only by the lags of other variables; hence, the  $e_t$  vector comprises composites of random shocks in the primitive system. In order to use the Johansen test, the above VAR needs to be converted into a VECM. The VECM is the appropriate model as it enables us to capture the long-run as well as the short-run dynamic relationships between the variables (Brooks, 2014). The VECM model is presented as follows:

$$\Delta Y_t = \delta_0 + \Pi Y_{t-1} + \sum_{i=1}^{n-1} \Gamma_{iy} \Delta Y_{t-1} + \varepsilon_t \quad (4.20)$$

where,  $y_t$  is a  $k \times 1$  vector of endogenous  $I(1)$  variables;  $\mu_0$  represents the intercept coefficients;  $\Pi$  is a  $k \times k$  long-run multiplier matrix; and  $\Gamma_{iy}$  are  $k \times k$  coefficient matrices describing the short-run dynamic effects.  $n$  is the lag length. Since the VECM representation is in differenced form, the lag order reduces to  $p-1$ .  $\varepsilon_t$  is a vector of independently and identically distributed innovations with a zero mean. In this study, the lag length selection criteria include the AIC, SBC, EFP and HQC as discussed under the section on lag length selection criteria which suggest that a second order ( $p=2$ ) VAR model is most appropriate.

The VAR model has pros and cons as any other econometrics estimation techniques. Firstly, it is simpler and the researcher does not need to identify which variable should be endogeneous or exogenous as mentioned earlier. The model itself treats all

variables as endogenous. Secondly, each VAR regression can be estimated by using simple OLS, (Asteriou & Hall, 2015). The VAR model has been criticised for a number of limitations, and it has been indicated that it is not easy to interpret the coefficient of multiple equations. However, the advocates of the VAR model introduced the use of impulse response functions which examine the effect of an unanticipated shock (captured in the error term) of each dependent variable in VAR model, the remaining variables over time.

The VECM approach has been adopted for its advantages: it measures the correction in deviations from equilibrium of the dependent variable relative to its covariates in the previous period; eliminates trends from variables; and solves the spurious regression problems. Thus, the ECM are formulated in terms of first-order difference (Asteriou & Hall, 2015). The disequilibrium in error terms is a stationary variable; therefore, the VECM have a significant effect because, when two variables are found to trend together indicating that there is some adjustment process, this prevents the error in the equilibrium relationship from becoming larger.

#### **4.9. POST-ESTIMATION DIAGNOSTIC ANALYSIS**

The characteristics of the best model need more attention to residuals diagnostic inspections to detect pathologies that might exist or affect the regression. Furthermore, closer inspection to regression pathologies helps the researcher ensure that none of the OLS assumptions are violated. For this purpose, section 4.9.1 gives full discussion of all regression pathologies together with relevant remedial measures that will be considered in the study to maintain accurateness and reliability of the model. Under each aspect, relevant tests to detect existence of regression problems are also highlighted. The first aspect discussed in this regard is heteroscedasticity, followed by autocorrelation, normality of residuals, multicollinearity and functional form. The last section which is 4.9.2 discusses model stability tests adopted in the study as a way to detect structural breaks that might exist along the sample period.

##### **4.9.1. Heteroscedasticity**

Among classical linear regression assumptions is that residuals should be homoscedastic which means that variance of the residuals must be equal or constant for the model to be good; therefore, failure to comply with this assumption would result

in an undesirable model. This violation of the assumption results in heteroscedasticity which is not good in a regression model. Although heteroscedasticity usually arises in cross-sectional data, it might affect time-series data at any time. The problem of heteroscedasticity mostly affects the Gauss-Markov theorem which claims that OLS estimators are Best Linear Unbiased Estimators (BLUE). Although the estimators remain unbiased and consistent provided the fact that residuals are not correlated with any of explanatory variables, the existence of heteroscedasticity in a model tends to make estimators to be inefficient. As BLUE assumptions reflect that efficient estimators are those with relatively small variances, in the case of unequal variances, it might not be as small as required by the assumption. Other than compromising the BLUE assumption, it usually leads to higher than anticipated t-stats and F-stats through underestimating variance of estimated parameters. A higher than usual t-statistic and F-statistic might mislead us to conclude that the coefficients are statistically significant even if it does not influence the dependent variable at all.

Relevant tests with regards to this problem are necessary in ensuring that OLS assumptions are not violated and to preclude inaccurate conclusions that might be effected by this regression pathology. Logarithmic transformation of all variables is the first option to avoid the problem. Heteroscedasticity can be tested through various tests, but in this study, the Breusch-Pagan (1979) LM test is used. The null hypothesis claiming that residuals are not heteroscedastic should be rejected if the computed p-value is below the 0.05% level of significance. The conclusion in this regard would be that residuals are heteroscedastic which would reveal inaccuracy of the model.

#### **4.9.2. Autocorrelation**

According to classical linear regression assumptions, residuals are not serially correlated. If the residuals are dependently distributed which means that residuals are serially correlated as opposed to this assumption, the model might be inaccurate. This problem usually arises as an outcome of following factors: omitted variables, misspecification of the model and systematic errors in measurement. In the existence of serial correlation, although OLS estimators of the parameters maintain their feature of being unbiased and consistent, they become inefficient given that the variance is no longer minimal as required. Therefore, the Gauss-Markov theorem of BLUE estimators has been violated. This would cause the conclusion to be unreliable as a

result of overestimated  $R^2$  showing a better fit of the model than truly exists and a higher than expected t-statistic emphasising a greater significance of coefficients than necessary. The most recognised test in attempt to inspect the presence of serial correlation is the DW test. It is given by:

$$d = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2}$$

The interval of (d) ranges between 0 and 4 as, according to the rule of thumb regarding this test, the value nearer to zero indicates the presence of positive serial correlation while any value above two and closer to four shows a negative serial correlation. The value in which residuals have no serial correlation is 2 or any number closer to this value, thus the null hypothesis cannot be rejected for both positive and negative autocorrelations. As the serial correlation is categorised into first and higher order, it is argued that the DW test cannot account for higher order serial correlation. This inability of the DW test implies the possibility of its producing inconclusive results and failure to apply in the case of lagged dependent variables. This has led many researchers to apply the Breusch-Godfrey (1978) LM test which satisfies all the drawbacks of the DW test. The conclusion in this regard is usually guided by the LM statistic which is given by  $(n-p) R^2$ , in that, if this value goes beyond the critical value, the null hypothesis should be rejected. Therefore, we can conclude that there is serial correlation.

#### **4.9.3. Normality**

A fundamental assumption of the classical linear regression model is that residuals are normally distributed with a constant variance and zero mean. This assumption is essential in ensuring the correctness of the formulae that we use in calculating the t-test and the f-test. This means that violation of this assumption would render these inferential statistics invalid which in turn might lead us to draw wrong conclusions in hypothesis testing. The most commonly applied test of normality is the JB test to detect violation of the assumption. Basically the JB test is based on skewness and kurtosis of residuals, which normally anticipate coefficients of zero and three respectively (Mutsau 2012). Any value other than these stipulated coefficients such as skewness

greater one and kurtosis of greater than three would mean that residuals are not symmetrical by the mean which also confirms non-normality. The null hypothesis which claims that residuals are normally distributed would be accepted if its p-value is greater than 0.05 level of significance and a relevant conclusion could be drawn. The conclusion where the null hypothesis is accepted also confirms unbiasedness, efficiency, and consistency of the estimates in the model as indicated by Gauss-Markov theorem.

#### **4.9.4. Multicollinearity**

Classical linear regression assumptions also require no linear relationship between sample values of dependent variables, if such a situation exists among explanatory variables, the regression pathology called multicollinearity arises. This problematic regression pathology is usually differentiated into perfect and imperfect multicollinearity depending on the nature of the linear dependency between explanatory variables. The existence of perfect multicollinearity among exogenous variables leads to the impossibility of any estimation method providing unique estimates of the population parameters, which means that OLS estimators under this pathology tend to be not as unique as anticipated. In most cases, estimation of equation that has multicollinearity would be rejected in any software program. Perfect multicollinearity generally is more problematic than imperfect multicollinearity, that is, in the presence of imperfect multicollinearity Biasness of estimators remains unaffected. However, the estimated coefficients might be rendered imprecise due to large standard errors and variances. Large standard errors basically could, in turn, lead to removal of influential variables in the regression as result of low t-stats which might fail to give affected coefficients sufficient statistical significance. Usually in attempts to analyse data to ensure that all explanatory variables are not linearly correlated, the variance inflation factor analysis is applied, but it is rarely used in economic research. Thus, there is no formal test to detect presence of this pathology except a pair-wise correlation coefficient which is regarded as sufficient to indicate whether the problem exists or not. It reveals existence of multicollinearity in the case of coefficient value of 0.9 or above (Asteriou & Hall, 2007).

According to Cameron (2005:274), multicollinearity could simply be removed through application of the following measures: extending the sample size or by using monthly

or quarterly frequencies rather than annual data; dropping off variables that are rendered to be statistically insignificant in the model; or re-specifying the model by excluding linear related variables and including nonlinear related variables to avoid the problem of this pathology. Given the probability of these remedial measures leading to misspecification bias in the model, this pathology should be neglected if it is imperfect because estimations can be precise under imperfect multicollinearity. Hence, imperfect multicollinearity in this study would be not taken into consideration.

#### **4.9.5. Functional form**

Ramsey (1969) developed a regression specification error test (RESET) that is directed at inspecting possible problems in the regression model caused by inclusion of insignificant explanatory variables or omission of most important variables in applying the model. It is argued that exclusion of significant explanatory variables can effectively cause violation of the assumption that residuals have a zero mean as those variables ultimately form part of the error terms. The possibility of misspecification of the regression model would be generated during the collection of data and the incorrect mathematical presentation form of the model (Mutsau, 2012). The RESET test has complementary F and LM forms which are applicable in testing misspecification bias. The null hypothesis of correct specification or no misspecification should be rejected if the p-value of computed F or LM values is less than a stipulated significance level. Under such occasions we conclude that the model is misspecified.

#### **4.10. CAUSALITY TEST**

In econometrics, causality refers to the ability of a single variable to explain the other. The features of the VAR approach are that it allows the researcher to test the direction of causality of each variable in the model. Suppose we have two variables (X & Y) which affect with previous values. The VAR model captures the relationship between these two variables. The causality among variables may indicate a uni-directional relationship: (i)  $Y_t$  to cause  $X_t$ , (ii)  $X_t$  to cause  $Y_t$ ; (iii) there may indicate a bidirectional feedback; and (iv) there may be independence between variables. The appropriate procedure should be followed that allows for testing and statistically indicating the effect and cause of the relationship among variables (Asteriou & Hall, 2015). Granger

(1969) developed a causality test and Sim (1972) later came up with an alternative approach for testing causality among variables. These two tests are discussed in detail below.

#### 4.10.1. The Granger Causality Test

It is important to examine the nature of short-run and long-run Granger Causality relationship in the VAR-VECM after the cointegration relationship has been confirmed between FLFP and its explanatory variables. The Granger Causality theorem indicates that the present is caused by the past (Takaendesa & Odhiambo, 2007). Asteriou and Hall (2015) postulate that variable  $Y_t$  is said to cause  $X_t$  if  $X_t$  can be estimated using the past values of  $Y_t$ . The series of  $X_t$  causes  $Y_t$ , if the lagged values of  $X$  contribute significantly to developing a more accurate forecast of  $Y$ . The opposite is true, if the previous values of  $Y_t$  are statistically different from zero and lagged  $X$ -values are insignificant, then  $Y_t$  causes  $X_t$ . In a case where there is bidirectional causality, both  $X_t$  and  $Y_t$  are statistically significant and different from zero whereas independence occurs when both variables  $X$  and  $Y$  series are insignificant (Dougherty, 2011; Mutsau, 2013) as indicated in the equations below.

The Granger Causality test for two stationary variables involves the estimation of causality of  $X$  on  $Y$  and vice versa through application of simultaneous and multivariate equations on the VAR regression as follows where both error terms are assumed to be uncorrelated white-noise:

$$Y_t = \alpha_0 + \sum_{i=1}^n \alpha_{1i} Y_{t-1} + \sum_{i=1}^n \beta_{1i} X_{t-1} + \varepsilon_t \quad (4.21)$$

$$X_t = \beta_0 + \sum_{i=1}^n \alpha_{2i} Y_{t-1} + \sum_{i=1}^n \beta_{2i} X_{t-1} + v_t \quad (4.22)$$

In the above equations,  $n$  is the number of lagged variables. The errors terms  $\varepsilon_t$  and  $v_t$  are white-noise satisfying all classical assumptions. The null hypothesis of non-causality is tested between any pair of the variables in the model. There are four possible outcomes from these tests. For instance, as applicable to this study, between FLFP, output and unemployment, (i) there is no causality i.e. there is causal independence; (ii) there is forward uni-directional causality; (iii) there is reverse uni-directional causality; or (iv), there is feedback or bidirectional causality (Gujarati & Porter, 2009).

The causality test involves the estimation of the VAR given the above equations, checking whether the coefficients are significant or not and apply a variable deletion test (Asteriou & Hall, 2015). Furthermore, the conclusion will be based on the results of direction of causality produced by the test. In a case of single equation, the following steps should be considered. Firstly, we estimate the restricted model of  $Y_t$  on the lagged values of  $Y_t$  and obtain  $RSS_r$  and we also estimate the unrestricted model  $Y_t$  for the lagged values of both  $X_t$  and  $Y_t$  and obtain  $RSS_u$ . The causality test is tested under the hypothesis which says that;

$$H_0: \sum_{i=1}^n \beta_i = 0 \text{ or } X_t \text{ does not cause } Y_t \text{ (null hypothesis)}$$

$$H_0: \sum_{i=1}^n \beta_i \neq 0 \text{ or } X_t \text{ does cause } Y_t \text{ (alternative)}$$

The F-statistic is calculated by adopting the normal Wald test,  $f = \frac{RSS_R - RSS_U/m}{RSS_U/(n-k)}$ . If the f values are greater than the critical values, then we reject the null hypothesis and say that  $X_t$  causes  $Y_t$ .

#### 4.10.2. The Sims Causality Test

Sims (1980) suggested an alternative causality test because it is impossible for the future to cause the present. Sims suggested the following VAR model to check whether  $Y_t$  is caused by  $X_t$  or is the vice versa.

$$Y_t = \sigma_1 + \sum_{i=1}^n \beta_i X_{t-i} + \sum_{j=1}^w \gamma_j Y_{t-j} + \sum_{\rho=1}^k \theta_\rho X_{t-\rho} + \varepsilon_{1t} \quad (4.23)$$

$$X_t = \sigma_2 + \sum_{i=1}^n \theta_i X_{t-i} + \sum_{j=1}^w \delta_j Y_{t-j} + \sum_{\rho=1}^k \vartheta_\rho Y_{t-\rho} + \varepsilon_{2t} \quad (4.24)$$

Asteriou and Hall (2015) postulate that apart from lagging the values of X & Y, this test indicates that there are also leading values of X (Eq 4.23) and Y (Eq 4.24). When examining Eq 4.24, if Y causes X, then it should be expected that some of leading values of X would have a relationship with Y. In that case, the test indicates that, instead of testing the lagged values of X just like the Granger Causality analysis, the researcher can test for  $\sum_{\rho=1}^k \theta_\rho = 0$ . Furthermore, the causality runs to X from Y and not vice versa when the restriction is rejected since the test assumes that the future does not cause the present. Firstly, to carry out the Sims Causality test, the model is estimated without a leading term (restricted model), then estimates the model in eq

4.24 (unrestricted model) and obtains the F-statistics as the Granger Causality test. This test includes the leading terms which increase the number of regressors in the model, leading to greater loss of degrees of freedom. The study adopted the Granger Causality test because of its advantages over the Sims Causality test.

#### 4.11. MODEL STABILITY

##### 4.11.1. Chow Test

One of the classical linear regression model assumptions is that all parameters across the sample are constant. With reference to this assumption, model stability is most important in any regression. The most commonly used test for structural change that might have been resulted from new policies implemented or sudden shocks in the economy throughout the sample period is the Chow (1960) test. According to its procedure, observations must be separated into two sub-samples. For instance, if the structural break occurred in period T, observations of the period before and after period T should be estimated separately. A pooled regression model of all observations should also be conducted. After regressing the three models, the residual sum of squares from all regressions should be compared to see if there are changes in variances and deterministic terms caused by structural breaks. This would help to determine whether a single regression is more efficient than two separate regressions. Thus, this test requires clear knowledge of the exact time at which structural change occurred, given that limited information of the period at which structural change occurred might become a huge constraint in applying this test to check model stability. Error terms in all regressions must satisfy all classical linear regression assumptions such as normal distributions in all residuals with the same variance. The null hypothesis which claims that parameters are stable along the dataset could be rejected if F-stats exceed F-critical value. The relevant conclusion that parameters are structural unstable would be drawn. The F-stats is given by:

$$F = \frac{RSS_n - (RSS_{n_1} + RSS_{n_2})/k}{(RSS_{n_1} + RSS_{n_2})/(n_1 + n_2 + 2k)} \quad (4.25)$$

where RSS is the sum of squared residuals. Most researchers employ the cumulative sum of recursive residuals (CUSUM) tests and the cumulative sum of squares of

recursive residuals (CUSUM-SQ) to check for model stability due to some drawbacks suspected to exist in the Chow test.

### 1.1.1. CUSUM and CUSUM-Sq Test

The CUSUM test developed by Brown, Durbin and Evans (1975) has been used frequently as an instrument to check both parameters and model stability. This emanates from its application which requires no knowledge of a specific period at which the structural change occurred. Therefore, the CUSUM test can be applied more easily than the Chow test. This test has the potential to detect prediction errors that are independently random and those errors, if present, just cancel each other out (Cameron, 2005). The CUSUM test is given by this equation:

$$W_r = \sum_{j=k+1}^t \hat{u}_t^{(j)} / \hat{\sigma}_t \quad (4.26)$$

Where  $U_i$  are the recursive residuals and  $\sigma_i$  is the standard error. This test plots the cumulative sum at 0.05% level of significance. The null hypothesis claiming that parameters are stable should be rejected if cumulative sum goes beyond the region between the two critical lines  $\pm 0.948 \left( \sqrt{(n-k)} + 1.896(r-k) / \sqrt{(n-k)} \right)$ . A conclusion indicating instability among parameters will be drawn. According to Mutsau (2013), proper application of this test effectively enables the researcher to detect the non-zero mean of the recursive residuals which results from shifts in the parameters of the model. However, it is argued that the test becomes powerful only when shifts are contributed by various parameter shifts and fails to detect non-zero means generated by a few parameter shifts. The CUSUM of squares should be applied to confirm shifts in parameters. CUSUM-SQ is given by:

$$WW_r = \sum_{j=k+1}^r (\hat{u}_t^{(j)})^2 / \sum_{j=k+1}^t (\hat{u}_t^{(j)})^2 \quad (4.27)$$

In the same sense, the null hypothesis ought to be rejected if the cumulative sum of squares goes outside the critical lines at the 5% level of significance. For model stability purposes, this study uses the CUSUM and CUSUM-SQ tests which are regarded as more effective than the Chow test. If the conclusion based on findings

given by the test shows presence of structural instability, dummy variables to address the problem will be included in the regression model.

#### 4.12. IRF AND FORECAST ERROR-VARIANCE DECOMPOSITION

The impact of shocks of each variable in the FLFP model is inspected through the use of Impulse Response Fractions (IRF) and Forecast Error Variance Decomposition (FEVD). Impulse responses indicate the behavioural pattern or direction of a shock to the endogenous variable. It also illustrates the reaction of the dependent variable (FLFP in this case) to a shock of each explanatory variable (GDP, GDP squared and unemployment) (Mutsau, 2013). The response of a dependent variable to a shock in a VAR system is easily understood through the use of IRF. A shock is caused by an innovation or error term in each model in a VAR system which affects the endogenous variable. Under the IRF, it is assumed that the error term should be uncorrelated and the model must be in a stable equilibrium, Mahlo (2014). Rossiter (2002) postulates that impulse response does not require orthogonalisation of shocks and the order of variables is invariant.

In a VAR model, any shock or impulse enters a system through the vector of error terms. Lutkepohl (2014) indicates that a non-zero component of residuals responds to an equivalent change in the dependent variable which also lead to changes in the other variables in the following period in the system. The additional effect of a single non-zero element of residuals can be expressed by the VAR model as a vector moving average (Lutkepohl, 2014).

$$y_t = B(L)^{-1}\mu_t = \Pi(L)\mu_t = \sum_{j=0}^{\infty} \pi_j \mu_{t-j} \quad (4.28)$$

Where the  $\Pi$  represents the symmetrical  $n \times n$  coefficient matrix, and  $\pi$  donates the responses to  $\mu_t$ . The above equation only holds if the variables are  $I(0)$  or if the VAR is stable (Mutsau, 2013). The graphical illustration of IRF (figure 5.2) is presented in this study to observe the effect of structural innovations to variables in the FLFP model.

Variance Decomposition measures the percentage of forecast error variation that is explained by other variables in the short-run dynamics and interaction. It also indicates how the dependent variable moves based on its own shocks versus the shocks of its explanatory variables (Mahlo, 2014). Furthermore, the VDC provides information

about the importance of a shock of each variable in the VAR model. In the VDC, the shock of a certain variable will first affect itself before affecting other variables in the system due to its dynamic mechanism in a VAR system. Most often, a forecast error variance of a time-series variable explains itself (Mutsau, 2013). Moreover, the properties of the forecast error in exploring the interrelationship among variables in a VAR system is better understood through the use of VDC.

#### 4.13. SINGLE EQUATIONS

Preliminary results (not reported) showed that the yearly data did not provide robust results when using the above VAR/VECM methodology, hence the study opted to employ the cointegrating single equation models to identify the long-run relationship between FLFP, GDP and Unemployment.

##### 4.13.1. Fully-Modified Ordinary Least Squares (FMOLS)

FMOLS is the cointegration technique by Philips and Hansen (1990) expecting that there is a presence of a single cointegrating vector and includes modifying OLS estimates of both long-run parameters and their related t-qualities to oblige for any bias related to autocorrelation and endogeneity issues existing in OLS residuals (Harris & Sollis, 2003). Therefore, the subsequent estimator is asymptotically reasonable and has completely effective ordinary asymptotical properties taking into account the application of standard Wald tests applying asymptotic chi-square ( $\chi^2$ ) factual deduction (Belke & Czudaj, 2010). Taking after Belke and Czudaj (2010), the vector process( $Y_t, X_t$ ) is as follows;

$$Y_t = X_t\beta + D_t\gamma_1 + \mu_{1t} \quad (4.29)$$

where  $Y_t$  is the (1) dependent variable and  $X_t$  denotes the stochastic explanatory variables ruled by  $X_t = \Gamma_{21}D_{1t} + \Gamma_{22}D_{2t} + \varepsilon_{2t}$  with  $\Delta\varepsilon_{2t} = \mu_{2t}$ . Moreover,  $D = (D_{1t}, D_{2t})$  are the deterministic trend independent variables and  $\mu_{1t}$  is the error term with zero mean and covariance. The FMOLS estimator is given by:

$$FMOLS = \begin{bmatrix} \beta \\ \hat{Y}_1 \end{bmatrix} = [\sum_{t=1}^T Z_t Z'_t]^{-1} \left[ \sum_{t=1}^T Z_t \hat{Y}_T - T \begin{bmatrix} Y_{12} \\ 0 \end{bmatrix} \right] \quad (4.30)$$

where  $Z_t = (X_t', D_t')$ .  $Y_t' = Y_t - \hat{\omega}_{12} \hat{\Omega}_{22}^{-1} \hat{\mu}_2$  denotes the transformed data and  $\gamma_{12} = \hat{\omega}_{12} \hat{\Omega}_{22}^{-1} \hat{\Lambda}_{22}$  denotes the estimated bias correction term with the long-run covariance matrices  $\hat{\Omega}$  and  $\hat{\Lambda}$  and their respective elements captured using  $\mu_t = (\hat{\mu}_{1t}', \hat{\mu}_{2t}')$ .

#### 4.13.2. Canonical Cointegrating Regression (CCR)

The CCR estimation technique by Park (1992), is a non-parametric approach for statistical interpretation in a cointegrated model including changes of the integrated procedures using stationary parts to represent long-run relationships among regressors and the error term. As indicated by Han (1996), this technique yields asymptotically productive estimators and  $\chi^2$  inference. The CCR estimator is characterized as;

$$\theta_{CCR} = \begin{bmatrix} \beta \\ \hat{\gamma}_1 \end{bmatrix} = [\sum_{t=1}^T Z_t Z_t']^{-1} \sum_{t=1}^T Z_t Y_t' \dots \dots \quad (4.31)$$

where  $Z_t = (X_t', D_t')$ .  $X_t' = X_t - (\sum_{t=1}^{t-1} \hat{\Lambda}_2) \hat{\mu}_t$  and  $Y_t' = Y_t - \left[ \sum_{t=1}^{t-1} \hat{\Lambda}_2 \hat{\beta} + \begin{bmatrix} 0 \\ \hat{\Omega}_{22}^{-1} \hat{\omega}_{21} \end{bmatrix} \right]' \mu_t$  denote the transformed data; the  $\hat{\beta}$  coefficients are estimates of cointegrating conditions applying static OLS;  $\hat{\Lambda}_2$  is the second segment of  $\hat{\Lambda}$ ; and  $\Sigma$  signifies the evaluated contemporaneous covariance matrix of the error terms.

#### 4.13.3 Dynamic Ordinary Least Squares Test (DOLS)

The DOLS test is another contrasting option to the EG approach recommended by Stock and Watson (1993), the test is a parametric strategy in which the first difference (lagged) terms are clearly evaluated (Saayman, 2010). It augments the cointegrating model by expanding extending innovations with leads, lags and contemporaneous estimations of the regressors ( $\Delta X_t$ ) (Saayman, 2010), to such an extent that the new cointegrating condition error term is orthogonal to the whole history of the stochastic regressor improvements (Belke & Czudaj, 2010). The DOLS results in a more effective test for cointegration and gives fair estimates of the long-run relationship (Harris, 1995; Harris & Sollis, 2003). As indicated by Belke and Czudaj (2010), this technique accepts that the additional lags of ( $\Delta X_t$ ) totally remove the long-run correlation between error terms  $\mu_{1t}$  and  $\mu_{2t}$  respectively

#### **4.14. CHAPTER SUMMARY**

The chapter outlined the nature and sources of both the quarterly and yearly data that will be used to investigate the relationship between FLFP, GDP and unemployment rate in Chapter 5. Furthermore, the modeling technique and statistical tests adopted in this study were described in detail. It outlined all the processes that are expected to be followed in addressing the research problem, which included stationarity tests, lag order specification, cointegration tests, VAR/VECM used to estimate the quarterly data as well as the single equation methods (FMOLS, DOLS and CCR) for the yearly data. Moreover, model stability and diagnostic tests were also described. All steps discussed in this chapter will be followed during estimation of regression models and interpretation of empirical findings of the study in Chapter 5. Therefore, interpretation of tests such as descriptive analysis, diagnostic inspections and model stability tests as important for accurateness of the model are given in the next chapter.

## CHAPTER 5: EMPIRICAL ANALYSIS AND INTERPRETATION OF RESULTS

### 5.1. INTRODUCTION

This chapter provides a presentation and discussion of the statistical and econometrics analyses which were outlined in the previous chapter. The empirical evaluation of research hypothesis and achieving the objectives of the study as indicated in Chapter 1 is done in this chapter. The multivariate estimation ( VAR-VECM) methodology and the single equations (FMOLS, DOLS, CCR) are adopted for the purpose of achieving the third objective of the study. All models are estimated according to objectives of the study. For multivariate equations, the study utilise the quarterly data and, for single equations, the yearly data.

This chapter is organised as follows: section 5.2 provides the model specification as adopted from the empirical literature from other studies conducted nationwide. The VAR-VECM is presented as guided by the theory in Chapter 4. Section 5.3 presents a full discussion of descriptive analysis and data transformation which include the issues related to data conversions, logarithmic transformations, measures of centrality, and dispersion. The covariance-correlation matrix is then presented to check for the probability of multicollinearity in the model.

Section 5.4 presents informal and formal stationarity testing through the unit-root hypothesis. Bivariate statistical investigations are done through the Augmented-Dickey-Fuller (ADF) and the Phillips-Perron (PP) tests to assess the integration order of all variables. The summaries and interpretation of these tests are reported in table form. The VAR lag length selection criteria results are given in section 5.5 for lag selection in the VAR model leading to the forecast error variance decompositions as well as the IRF. Lag order selection is also pivotal in the pair-wise Granger Causality tests.

### 5.2. MODEL SPECIFICATION

The study investigated the relationship between FLFP, economic growth and unemployment, adopting the same procedure applied by Lechman and Kaur (2015) and Olivetti (2013). The long-run model estimated in the study is presented as follows where FLFP is expressed as a function of output and unemployment.

$$FLFP_t = f(GDP_t, UNEMP_t) \quad (5.1)$$

To complete the analysis, the study employed three distinct variables. Firstly we take FLFP as a proxy for women participating in the labour force which explains the share of women (15-64 years) of a country's total labour force (Barker, 2013). Second is the level of output that a country possesses. In this case, GDP is used as a proxy for output. GDP reflects the level of production in the economy and it is transformed to natural logarithmic form. Lastly, we introduce the level of unemployment which is expected to have an impact on the participation of women in the labour force in most developing economies.

$$FLFP_t = \alpha_0 + \delta_1 GDP_t + \delta_3 UNEMP_t + \mu_t \quad (5.2)$$

The subscript t represents time and  $\alpha_0$  is a constant while  $\delta_1, \delta_2$ , and  $\delta_3$  are long-run parameters. The error term,  $\mu$ , is a Gaussian white-noise process satisfying all the basic assumptions of the classical linear regression model. Output has been transformed to logarithmic form before taking its square and incorporating it into equation. From empirical literature, FLFP and economic growth exhibits a U-shaped relationship (Lechman & Kaur, 2015; Lincove, 2008; Tsani et al., 2012) where the participation of women is the response variable, and output and unemployment are the explanatory variables. The study analyses the FLFP (Eq 5.2) model specified above within a VAR/VECM framework and single equations.

The VAR-VECM of the FLFP model for South Africa can be expressed as follows;

$$y_t = \mu + \Gamma_1 y_{t-1} + \Gamma_2 y_{t-2} + e_t \quad (5.3)$$

The above equation is a second order VAR where  $y_t = (flfp_t, gdp_t, \& unemp_t)$  and is also a  $(3 \times 1)$  column vector,  $\mu$  indicates a  $(3 \times 1)$  constant vector, and  $\Gamma_1$  is a  $(3 \times 3)$  matrix of autoregressive coefficients.

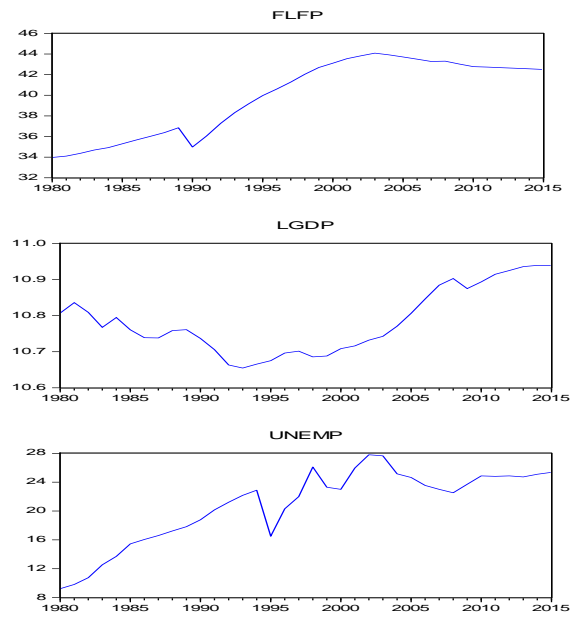
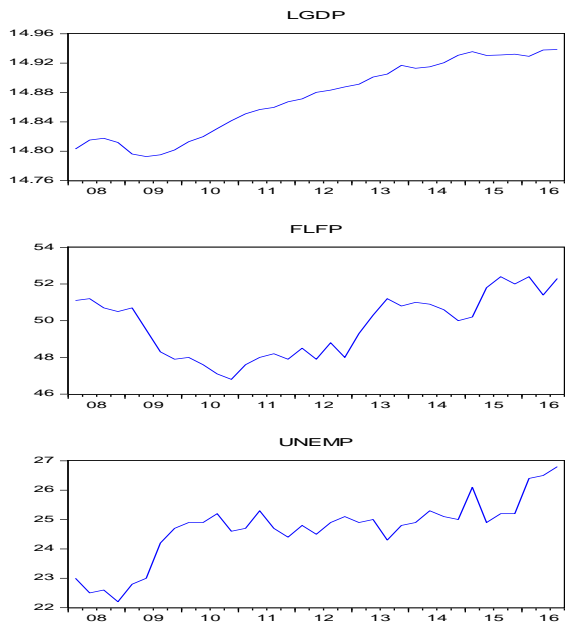
### 5.3. DESCRIPTION OF DATA

#### 5.3.1. Graphical Analysis of Data

Quarterly time-series variables are presented below. The positive trend in GDP is quite pronounced from 2009 Q1<sup>2</sup>, while the trend in FLFP switches from negative to positive from 2010 Q1 and unemployment also exhibits a mildly positive trend for most of the period, with a steep initial jump in 2008 Q3 and Yearly time-series variables are presented below, FLFP indicates a positive trend from 1990. Upon a cursory inspection, the data is likely to be non-stationary; however, rigorous statistical tests were conducted to confirm this hunch.

**Figure 5.1a: Graphical analysis of quarterly data (2008Q1 -2016Q3)**

**Figure 5.1b: Graphical Analysis of yearly data (1980 – 2015)**




---

<sup>2</sup> Q = quarter

Figure 5.2a: Graphical analysis of quarterly data at first difference (2008Q1 -2016Q3)

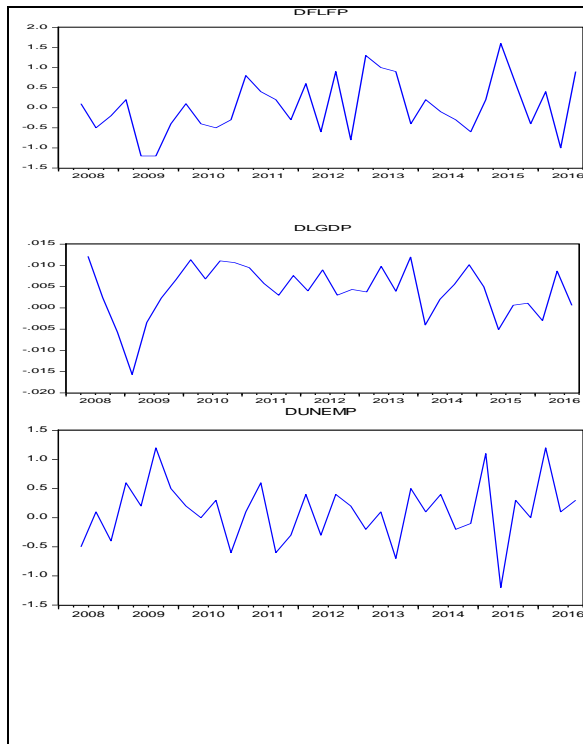
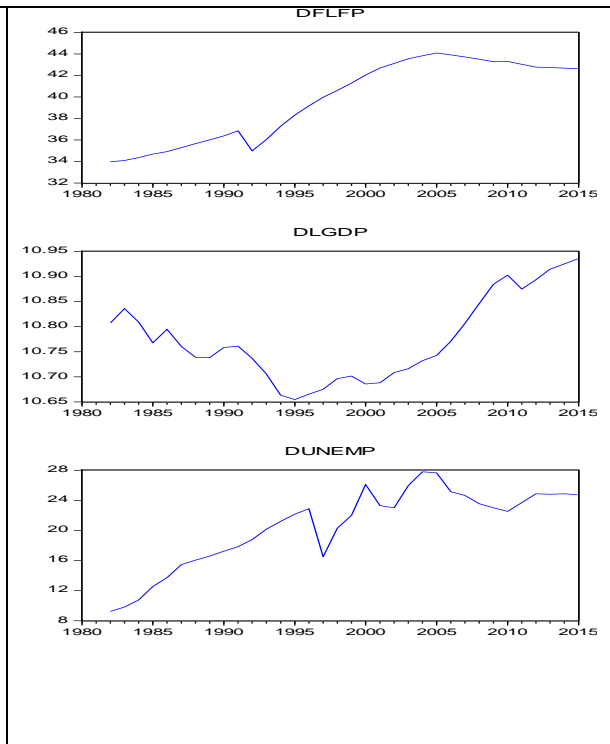


Figure 5.2b: Graphical analysis of quarterly data at first difference (1980 - 2015)



### 5.3.2. Descriptive Statistics

Table 5.1: Descriptive statistics: Results for quarterly series (2008Q1 -2016Q3)

	Mean	Standard-Dev	Minimum	Maximum	Obs.	Countries
<b>FLFP</b>	49.740	1.696	52.400	25.400	35	1
<b>LGDP</b>	14.872	0.050	14.703	14.938	35	1
<b>UNEMP</b>	24.668	1.086	22.200	26.800	35	1
	Skewness	Kurtosis	Jarque-Ber	Probability	Obs.	Countries
<b>FLFP</b>	-0.070	1.665	2.624	0.269	35	1
<b>LGDP</b>	-0.190	1.556	3.249	0.196	35	1
<b>UNEMP</b>	-0.578	3.252	2.069	0.355	35	1

Source: Author's own calculation

For this purpose, the summary of measures of centrality and dispersion, kurtosis and skewness are provided in Table 5.1. Measures of centrality are therefore computed with variables after logarithmic transformations. All variables (FLFP, LGDP and UNEMP) indicate that they are skewed to the left (negative sign); this is shown by a negative skewness coefficient. The kurtosis of a normal distribution is 3, and any measurement less than 3, indicates that a distribution is flatter compared to a normally

distribution, (Ziramba, 2007). The FLFP and LGDP kurtosis values are close to 2; hence, they are flatter than the theoretical normal distribution. A perfectly normally distribution is shown by unemployment with a coefficient of 3.2. The JB statistic which measures normality of variables is reported in this table with respect to its probability values under the null hypothesis that variables are normally distributed. According to the results, we fail to reject the null hypothesis and conclude that all variables are normally distributed, the probability values being greater than 0.05%. This provides evidence that outliers are not a problem in each individual sample of the data set. As part of descriptive statistics, a covariance-correlation matrix of all variables in the dataset was computed to understand the co-movement of variables against each other in Table 5.2 below.

**Table 5.2: Covariance-Correlation Matrix**

	FLP	LGDP	UNEMPL
FLP	[ <b>1.000</b> ] 2.796 -----		
LGDP	[ <b>0.514</b> ] 0.043 (0.0016)	[ <b>1.000</b> ] 0.002 ----	
UNEMP	[ <b>0.068</b> ] 0.122 (0.695)	[ <b>0.731</b> ] 0.039 (0.000)	[ <b>1.000</b> ] 1.146 -----

Source: Author's own calculation<sup>3</sup>

Basically, co-movement of variables as given by the covariance enabled the researcher to establish *a priori* expectations before estimation of regressions. The pairwise correlation assisted in measuring the presence of multicollinearity and the researcher could use it to set *a priori* expectations of all variables before full estimation of the model. It is argued that a correlation above 0.9 reflects presence of multicollinearity among respective variables (Mutsau, 2012). In Table 5.2 covariance, correlation and probability are presented chronologically with correlation in square

<sup>3</sup> Notes: Bold figures in square brackets represents the correlation coefficients, while italicised figures represents variances or covariances and figures in rounded brackets represent probabilities

brackets and probability values in normal brackets. According to the statistics shown by the covariance, all variables are moving together. The pair-wise correlation displays a positive linear relationship, and the multicollinearity is not a problematic issue among variables.

### **5.3.3. Correlation Matrix**

With reference to Table 5.2, the correlation between the LGDP and FLFP is 0.51, while LGDP and unemployment correlation is 0.73 and the correlation for FLFP and unemployment is 0.07. All the correlations are below the 0.8 rule of thumb; hence, one may conclude that multicollinearity is not a problem.

## **5.4. STATIONARITY TESTING**

The importance of establishing the order of integration through informal and formal methods follows the assertion by Engle and Granger (1987) that more than one time-series variables are found to be non-stationary,  $I(1)$  individually, but a linear combination of them is found to be  $I(0)$  (ie., their error terms from a regression are stationary, in which case, the time-series linear combination has a long-run relationship; hence, the variables are cointegrated. Formal and informal methods were conducted in this study to confirm order of integration of all variables, this significantly would allow us to compute cointegration analysis. The informal method included graphical representation of all variables in the levels indicated in the above section. All time-series through graphical illustration reflected upward trends which shows that they are not mean-reverting, but after first-differencing, trends disappeared which leads to the conclusion that variables are integrated of order one  $I(1)$ . Correlograms also form part of informal method of testing for stationarity although not applied in this study. Formal methods were also applied because conclusions with respect to integration property of variables could not be based on graphical analysis only. There are several tests that can be used as formal methods to establish whether a time-series variable contains a unit-root or not; however, in this study only the ADF and Phillips-Perron (PP) tests were applied, where the PP test was used to confirm the ADF test results. Under the null hypothesis that a time-series has a unit-root, we ought to reject it when the t-statistic is more negative compared to critical values at conventional significance level (1%, 5%, 10%) and it may not be rejected if the t-

statistic is less than critical values. The critical values are obtained from the MacKinnon (1996) tables.

Table 5.3 below presents the summary of these tests with the ADF using the Schwarz Information Criterion for selection of optimal lag lengths while the bandwidths for the PP-tests are indicated using the Newey-West Bartlett Kernel selection as in other empirical works. The asterisks \*, \*\* and \*\*\* indicate statistical significance at 1%, 5% and 10% respectively. All variables are tested for intercept, intercept & trend and neither intercept nor trend as discussed above in Chapter 4. Both tests revealed that all variables are non-stationary in their levels which means that we have evidence to accept the null hypothesis that the time-series in consideration has a unit-root. However, after first differencing, all of them become stationary, thus variables are integrated of order one, I(1). Hence, we would follow the Johansen Maximum Likelihood cointegration approach to conduct the cointegration test given that variables are integrated of order one, I(1).

**Table 5.3: Summary of unit-roots results**

Variable	Model	ADF test			Critical Values	PP test		Critical Values	
		Lag	$\tau_\tau, \tau_\mu, \tau$			BW			
FLFP	Intercept			1%	-3.639			1%	-3.639
		0	-0.853	5%	-2.951	0	-0.845	5%	-2.951
				10%	-2.614			10%	-2.614
	Intercept & trend			1%	-4.252			1%	-4.252
		0	-1.972	5%	-3.548	8	-1.919	5%	-3.548
				10%	-3.207			10%	-3.207
None			1%	-2.634			1%	2.634	
	0	0.265	5%	-1.951	1	0.259	5%	-1.951	
			10%	-1.610			10%	-1.610	
DFLFP	Intercept			1%	-3.646			1%	-3.646
		0	-5.220**	5%	-2.954	2	-5.250	5%	-2.954
				10%	-2.615			10%	-2.615
	Intercept & trend			1%	-4.262			1%	-4.262
		0	-5.701	5%	-3.552	8	-5.809	5%	-3.552
				10%	-3.209			10%	-3.209
None			1%	-2.636			1%	-2.636	
	0	-5.294	5%	-1.951	2	-5.324	5%	-1.951	
			10%	-1.610			10%	-1.610	

Variable	Model	ADF test	Critical Values	PP test	Critical Values	
<b>Conclusion: non-stationary, integrated of order one I(1)</b>						
LGDP	Intercept		1% -3.639		1% -3.639	
		1	0.451	5% -2.951	1 -0.498	5% -2.951
				10% -2.614		10% -2.614
	Intercept		1% -4.252		1% -4.252	
	& trend	1	-1.396	5% -3.548	1 -1.671	5% -3.548
				10% -3.207		10% -3.207
DLGDP	Intercept		1% -2.634		1% -2.634	
	None	1	3.794	5% -1.951	1 3.243	5% -1.951
				10% -1.610		10% -1.610
	Intercept		1% -3.646		1% -3.646	
		0	<b>-3.870</b>	5% -2.954	0 -3.870	5% -2.954
				10% -2.615		10% -2.615
UNEMP	Intercept		1% -4.262		1% -4.262	
	& trend	0	-3.813	5% -3.552	1 -3.834	5% -3.552
				10% -3.209		10% -3.209
	Intercept		1% -2.639		1% -2.636	
	None	0	-2.329	5% -1.951	2 -3.363	5% -1.951
				10% -1.616		10% -1.610
<b>Conclusion: non-stationary, integrated of order one I(1)</b>						
UNEMP	Intercept		1% -3.639		1% -3.639	
		0	-1.175	5% -2.951	3 -0.971	5% -2.951
				10% -2.614		10% -2.614
	Intercept		1% -4.252		1% -4.252	
	& trend	0	-2.418	5% -3.549	1 -2.330	5% -3.549
				10% -3.207		10% -3.207
DUNEMP	Intercept		1% -2.636		1% -2.636	
	None	1	1.699	5% -1.951	4 1.500	5% -1.951
				10% -1.611		10% -1.611
	Intercept		1% -3.646		1% -3.646	
		0	-7.500	5% -2.954	2 -5.250	5% -2.954
				10% -2.615		10% -2.615
DUNEMP	Intercept		1% -4.262		1% -4.262	
	& trend	0	-8.190	5% -3.552	8 -5.809	5% -3.552
				10% -3.209		10% -3.209
			-7.502	1% -2.636		1% -2.636

Variable	Model	ADF test	Critical Values		PP test	Critical Values		
	None	0	5%	-1.951	2	-5.324	5%	-1.951
			10%	-1.611			10%	-1.611

**Conclusion: non-stationary, integrated of order one I(1)**

*Source: Researcher's estimation*

## 5.5. THE VAR AND VECM ESTIMATION PROCESS

This section of the study covers the interpretation of the estimated results of the vector autoregressive and VECM, involving quarterly data from 2008 Q1 to 2016 Q3.

### 5.5.1. Lag Length Selection

When applying the VAR modeling leading to Granger Causality testing, IRF and variance decompositions, it is common practice to select the optimal lag length through the lag length selection criterion.

The information criteria applied in Table A1, Appendix A, are the sequential modified Log-Likelihood ratio tests (LR), the FPE, the Schwarz Bayesian (SBC) criterion, the Akaike (AIC), and the Hannan-Quinn (HQ) criteria as discussed in Chapter 4. Table A1, Appendix A, presents the lag length results indicating the optimal lag length that should be selected in modeling the VAR. All the approaches suggest a lag length of 1 ought to be selected except for the AIC which selects a lag length of 2. The study estimated both 1<sup>st</sup> and 2<sup>nd</sup> order VAR models; however, the second order VAR model provided superior results. Thus, the study pursued a thorough investigation of a 2<sup>nd</sup> order VAR. This choice was further reinforced by the AIC criterion which performs well in the context of a small sample (see De-Graft Acquah, 2010, for example).

### 5.5.2. Stability of the VAR

The study adopted the Autoregressive (AR) roots test to examine the stability of the VAR process and found that roots lay outside the unit circle; thus the VAR is unstable, as shown in Table B2 and Figure B2 (Appendix B). Johnston and DiNardo (1997) postulate that if each root has a modulus equal to 1 or closer to 1, all variables in the VAR system are I(1) and therefore the variables to be estimated in the VAR model require differencing. Since all the moduli in the AR table are strictly close to 1 and marginally greater than 1, a VAR estimate may be too unstable to appropriately

capture short-run interactions in the dynamic model of this study. This is evidence for the need to test for cointegrating relations and for short-run adjustment dynamics relative to these potential long-run cointegrating relationships. Given the results of the AR roots test, both the VAR and VECM are plausible models to consider; hence, the approach this study takes is to estimate both the VAR and VECM in order to address the objectives and hypotheses of the study.

### **5.5.3. VAR Estimates**

This section of the study presents the long-run relationship and causality among variables under the VAR system, following the ordering of FLFP-LGDP-UNEMP. The study tested whether the specified ordering of variables would produce implausible and insignificant results for some variables. Moreover, the unit-root discussed above indicates that the VAR model is unstable (see Appendix A, Table A2 & Figure A2). The results of the estimated VAR are presented in Appendix A (Table A3). Impulse response and variance decomposition are also presented in section 5.3.3.1 & 5.3.3.2 under the estimated VAR model. Granger Causality test is presented based on the same VAR model in section 5.4.5 below.

#### **5.5.3.1. Impulse Response Function**

In this section the structural model of the errors of the VAR is analysed through IRF graphs below. The IRF presented below show the economic relationship between FLFP and its explanatory variables (economic growth and unemployment). Thus, it shows the evolution of economic shocks in the VAR system.

**Figure 5.2: Response to Cholesky One Standard Deviation Innovation**

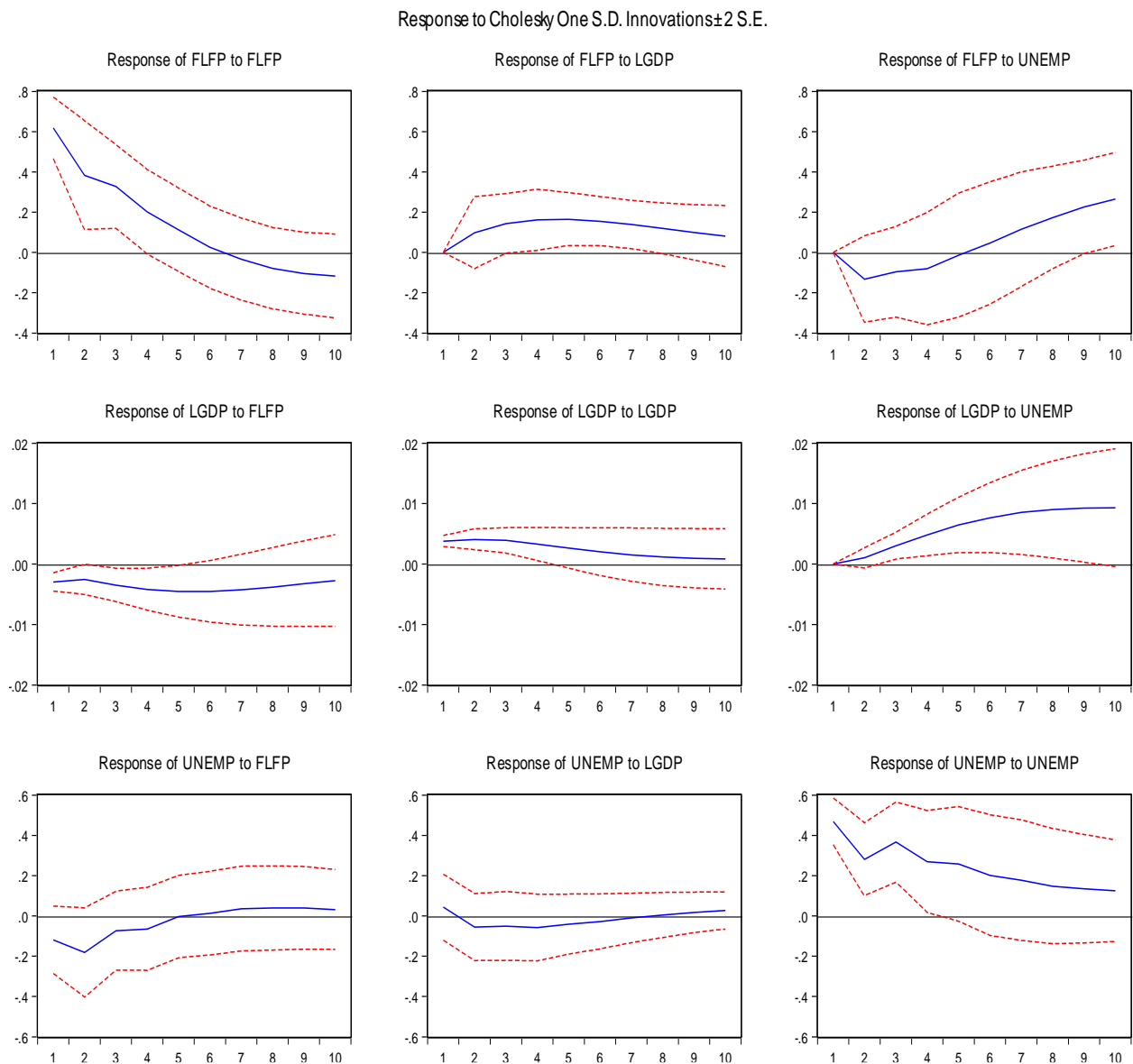


Figure 5.2 above indicates the response of the participation of female in the labour force in the VAR system when a shock is put to the error term. A unit shock was applied to each variable to see its effect on the system. To moderate the problem of contemporaneous correlations in the residuals of the VAR system, the Cholesky decomposition degree of freedom adjusted was chosen. The ordering of variables in the Cholesky decomposition is guided by the theoretical relationship between FLFP and its determinants. The function is estimated by Cholesky ordering as FLFP, GDP and unemployment respectively. The Cholesky decomposition or recursive ordering scheme for this study implies that the first variable (FLFP) is not correlated with any other variable of the VAR; the second variable (GDP) is contemporaneously correlated

only with the preceding FLFP variable; while last variable (UNEMP) is correlated with all the preceding variables.

FLFP is sensitive to a unit standard deviation shock in the error terms associated with itself, GDP and UNEMP. The first row shows the response of FLFP to its own shock; thereafter, the second and third columns of the first row capture the response of LGDP and UNEMP, respectively. First, FLFP responds immediately to a unit shock to itself with a rise in female employment by 0.3% to an unanticipated 1%; however, by the 5<sup>th</sup> quarter equilibrium is restored. Third, the IRF of the first row, the second column shows the response of FLFP to a shock to GDP, an unanticipated 1% in GDP causes the female participation rate to rise gradually. By the third quarter, the rise in FLFP is almost 0.2%. The FLFP returns slowly over the long-term (in excess of 10 quarters) back to equilibrium. The response of FLFP to a unit standard deviation shock to UNEMP is an initial decline in FLFP, a 1% shock to UNEMP results in a 0.1% decline in the first 1.5 quarters but thereafter, FLFP continues to rise, By the fifth quarter, its equilibrium level is exceeded while it continues indefinitely on an upward trajectory.

Next the IRF results depicted in the first column of rows two and three are relevant to this study in assessing the dynamic relations between FLFP and LGDP and UNEMP, respectively. First, the natural log of GDP appears nonresponsive to a unitary standard deviation shock to FLFP; a 1% unanticipated shock appears to marginally cause LGDP to fall below equilibrium, and a 1% unanticipated rise in FLFP causes GDP to fall by 0.01%; hence, one may conclude that GDP is strongly inelastic to a rise in FLFP. Second, the UNEMP falls below equilibrium to a low of about 0.2% after 1.5 quarters to a 1% unanticipated shock to FLFP but thereafter retraces back to its equilibrium position by the fifth quarter. Here again, one may conclude the response of UNEMP is relatively inelastic.

The above results show that some effect is reliable with economic theory; however, the magnitudes appear to be insignificant. The results suggest that the participation of women in the labour force in the South African does not provide significant short-run results but is more likely to be a long-term phenomenon. However, the limitation of this study is that the quarterly dataset only spans about 8 years which is too short for meaningful short-run relations to emerge in the analyses.

### **5.5.3.2. Variance decomposition**

The ten quarterly variance decomposition results of the VAR model are presented in Appendix A in Tables A5 to A7 respectively. Table A5 shows that FLFP is mainly driven by its own shock, given 100% in the first quarter. The decline began in the second quarter throughout the forecast period from 95.14% to 66.89%. However, shocks in FLFP seem to explain the behaviour of FLFP compared to LGDP and UNEMP. The contribution of LGDP and UNEMP appear to be insignificant, ranging from 0.00 (first period) to 14.57 and 18.53% (last period) in the forecast.

Table A6 indicates that economic growth is largely affected by its own impulses and the impulses of the participation of women in the labour force in the first year, accounting for 61.78% and 38.22% of the variations respectively. The contribution of LGDP gradually declines to 10.97% while the contribution of FLFP decreases at a constant rate to 20.34% at the end of the forecast period. More interestingly, the increasing shocks to UNEMP account for changes in economic growth with a contribution of 68.69% at the tenth quarter. These actually indicate that it plays a significant role in explaining fluctuation in the latter over time.

Table A7 reports that innovations in FLFP and LGDP are insignificant to explain movement in UNEMP in the first quarter of the forecast period, given their percentage values of 6.01 and 0.80 respectively. From the second period, the contribution of economic growth stays relatively constant at about 1.4% to 1.8% while the contribution of FLFP from 6% to 13.46%, then declines in the third period to 8.04% by the end of the forecast period. Notably, 93.19% of the variance in unemployment is due to its own shocks in the first quarter of the forecast period. In the second period of the forecast, unemployment responds negatively on its own shock. After the third period, it rises from 87.79% to 90.09% throughout the forecast period. This implies that the variance decomposition plays a significant role in explaining fluctuation in the latter over time.

### **5.5.3.3. VAR diagnostic test results**

The VAR diagnostic tests were conducted to examine the model's statistical reliability and adequacy. The results are displayed in Appendix A, Table A4. FLFP is driven by its own shock as indicated by 100% in the first quarter.

i) Serial correlation

By doing the autocorrelation test and setting the null hypothesis,  $H_0$ : No autocorrelation. The serial correlation LM stats indicate the p-values greater than 5% at a selected lag length. However, we fail to reject the null hypothesis and conclude that there is no serial correlation.

ii) Heteroscedasticity

To evaluate whether the fourth classical linear regression model (CLRM) assumption is violated or not, we test for heteroscedasticity. By setting the null hypothesis;  $H_0$ : Homoscedasticity, the heteroscedasticity point is 71.16 with p-value of 0.50 which is greater than 5%. Therefore, we fail to reject the null hypothesis and conclude that heteroscedasticity does not exist; alternatively, we can say homoscedasticity exists.

iii) Normality test

The CLRM assumes that residuals are normally distributed. To test the validity of this assumption, we apply the JB test. The null hypothesis is that residuals are normally distributed ( $H_0$ : normally distributed). The p-value of the JB test jointly is 0.99, greater than 5% level of significance. Therefore, we fail to reject the null hypothesis and conclude that they are normally distributed.

#### **5.5.4. VECM Estimates**

As noted earlier, AR roots test the stability of the VAR process. It was found that the roots lie outside the unit circle, and the VAR is thus unstable, as shown in Table B2 and Figure B2, Appendix B. Hence, this subsection presents the results of the Johansen (1991) VECM analyses which take into consideration the long-run cointegrating relationship as well as the short-run adjustment relationships.

##### **5.5.4.1. Deterministic component**

Asteriou and Hall (2007; 2015) and Harris (1995) suggest the following five steps to determine whether an intercept and trend should be included in the VAR-VECM model. The cases to consider are as follows:

- Case 5: quadratic trend and intercept in the cointegrating equation, intercept and linear trend in the VAR system. This case is not a reasonable option since it is

problematic to interpret it from an economic point of view, unless there is good theoretical support to include this case.

- Case 4: both the cointegrating equation and VAR models have an intercept. The cointegrating equation shows a linear trend but not the VAR model. This case indicates that the model does not include the quadratic trends.
- Case 3: intercepts in the data with no trend in the regressions. It is assumed that the intercept in the cointegrating equation is cancelled out by the intercept in the VAR, therefore leaving only one intercept in the short run (Vuyeka, 2015).
- Case 2: no trends in cointegrating vector or in VAR model but an intercept in the VAR model.
- Case 1: No deterministic components in the cointegrating equation(s) or in the VAR. in the cointegrating equation(s). All intercepts are zero. This case is rarely applicable since the intercept is needed in order to account for adjustment in the unit of measurement of the variables in the VAR system.

**Table 5.4: Summary of cointegration test assumptions**

<b>Data Trend:</b>	<b>None</b>	<b>None</b>	<b>Linear</b>	<b>Linear</b>	<b>Quadratic</b>
Test Type	No Intercept	Intercept	Intercept	Intercept	Intercept
	No Trend	No Trend	No Trend	Trend	Trend
Trace	1	1	1	2	3
Max-Eig	1	1	1	2	3

\*Critical values based on MacKinnon-Haug-Michelis (1999)  
Source: Eviews

In the context of the above table, since cases 1 and 5 are deemed implausible for macroeconomic time-series data in practice, the researcher focuses on the remaining cases (2, 3 & 4). These three cases indicate a strong long-run relationship in the model. In case 2, when there is no restrictive condition, both the trace and max-eigenvalue tests confirm the existence of one cointegrating vector, as for case 3. Two cointegrating vectors are indicated by both tests in case 4 which allows for both a linear intercept and a trend in the cointegrating equation. Cases 2, 3 and 4 were estimated (not reported in the interest of brevity); however, case 2 provided the most robust results. Hence, the study reports on the results of the Johansen cointegration test and VECM estimation based on the case 2 underlying assumptions.

### 5.5.4.2. The Johansen Cointegration Test

Table 5.5 below reports the cointegration test results. The hypothesis of no cointegrating vector is rejected by both the trace and the maximum eigenvalue tests at 5% significance level. Their test statistics of 44.844 and 27.644 are greater than their respective critical values of 35.192 and 22.299. Moreover, the alternative hypothesis of the existence of a single cointegrating relationship is accepted since both tests indicate that their statistics are less than their critical values, 17.200 smaller than 20.261 (trace test) and the 9.841 is less than 15.892 (maximum eigenvalue). Hence, the analysis concludes that one long-run cointegrating equation exists among FLFP, LGDP and UNEMP. However, it supports the use of the error correction model (ECM) to show short-run dynamics.

**Table 5.5: Johansen Cointegration Test Results**

H <sub>0</sub>	Trace stats	Critical values (5%)	Max-Eugen	Critical values (5%)
None	44.84	35.19	27.64	22.29
At most 1	17.20	20.26	9.84	15.89
At most 2	7.36	9.16	7.36	9.16

Source: Own estimate

### 5.5.4.3. VECM results

The cointegrating vectors in the study have been confirmed by the above Johansen cointegration test, which allows the researcher to estimate the VECM which incorporates both the long run and short-run equilibrium relationships among variables. The three variable-VECM results are presented on Appendix A (see Table A7). The significance of each coefficient is determined by observing their probability values estimated through the OLS and reported in Table A8 (Appendix A) and also by comparing t-statistics against critical values. Each variable in the model is expressed in the VECM form as discussed in the previous chapter. Thus, the FLFP VECM model is basically expressed as follows:

$$\Delta FLFP_t = -0.05(-78.17 + FLFP_{t-1} + 6.28LGDP_{t-1} - 2.72UNEMP_{t-1}) - 0.005\Delta FLFP_{t-1} + 13.77\Delta LGDP_{t-1} - 0.38\Delta UNEMP_{t-1} \quad (5.1)$$

The estimated long-run coefficients for all the variables in the above equation  $(-78.17 + FLFP_{t-1} + 6.28LGDP_{t-1} - 2.72UNEMP_{t-1})$  are reasonable and have negative signs due to the error correction mechanism, however, they are interpreted as positive elasticities. Economic growth is negatively correlated to FLFP in the long run with a coefficient of 6.28. A 1% rise in economic growth has an insignificant effect on the participation of woman in the labour force. According to the above equation, the long-run unemployment coefficient is consistent with the economic theory and significant when the t-statistic is compared to the critical value. It implies that with a 1% increase in UNEMP, the labour force will increase by a magnitude of 2.7. This result is in line with what Barker (2013) found, and hypothesis 2 fails to be rejected.

In the short run, the FLFP adjustment coefficient of the cointegrating equation is of the correct size (-0.05) but is not statistically significant. It does not indicate any changes in the short run. If the coefficient indicated significance, it would have indicated that on average, 5% of the disequilibrium will be corrected quarterly when FLFP rises above its long-run relationship with the other variables, *ceteris paribus*. The insignificance of causality is understandable from the economic perspective but all these variables are more likely to have a significant impact on the economy in the long run. The correlation coefficients of the variables (13.77 & -0.38) are too statistically insignificant to influence the participation of women in the short run. Hence, the last objective of the study is achieved with a conclusion that a short-run relationship among variables does not exist. The LGDP VECM equation is presented as follows:

$$\Delta LGDP_t = -0.0012(-78.17 + FLFP_{t-1} + 6.28LGDP_{t-1} - 2.72UNEMP_{t-1}) - 0.0011\Delta FLFP_{t-1} + 0.15\Delta LGDP_{t-1} - 0.002\Delta UNEMP_{t-1} \quad (5.2)$$

The equation above treats economic growth as the dependent variable. Economic growth moves by 0.12% in the following period to cater for any increase in FLFP and UNEMP above its long-run equilibrium in the previous period. The speed of adjustment is highly statistically significant with a probability value of 0.0008%. However, due to empirical postulation, FLFP must decrease if UNEMP increases in order for the equilibrium to be restored. In other words, a rise in LGDP above its long-run structural average by 1% in the previous period, moves downward by about 0.12% in the following period.

The short-run coefficients of equation (5.2) indicate that neither of the variables is statistically significant. In the short run, economic growth does not have any significant relationship with any of the explanatory variables.

$$\Delta UNEMP_t = -0.008(-78.17 + FLFP_{t-1} + 6.28LGDP_{t-1} - 2.72UNEMP_{t-1}) - 0.26\Delta FLFP_{t-1} - 7.21\Delta LGDP_{t-1} - 0.31\Delta UNEMP_{t-1} \quad (5.3)$$

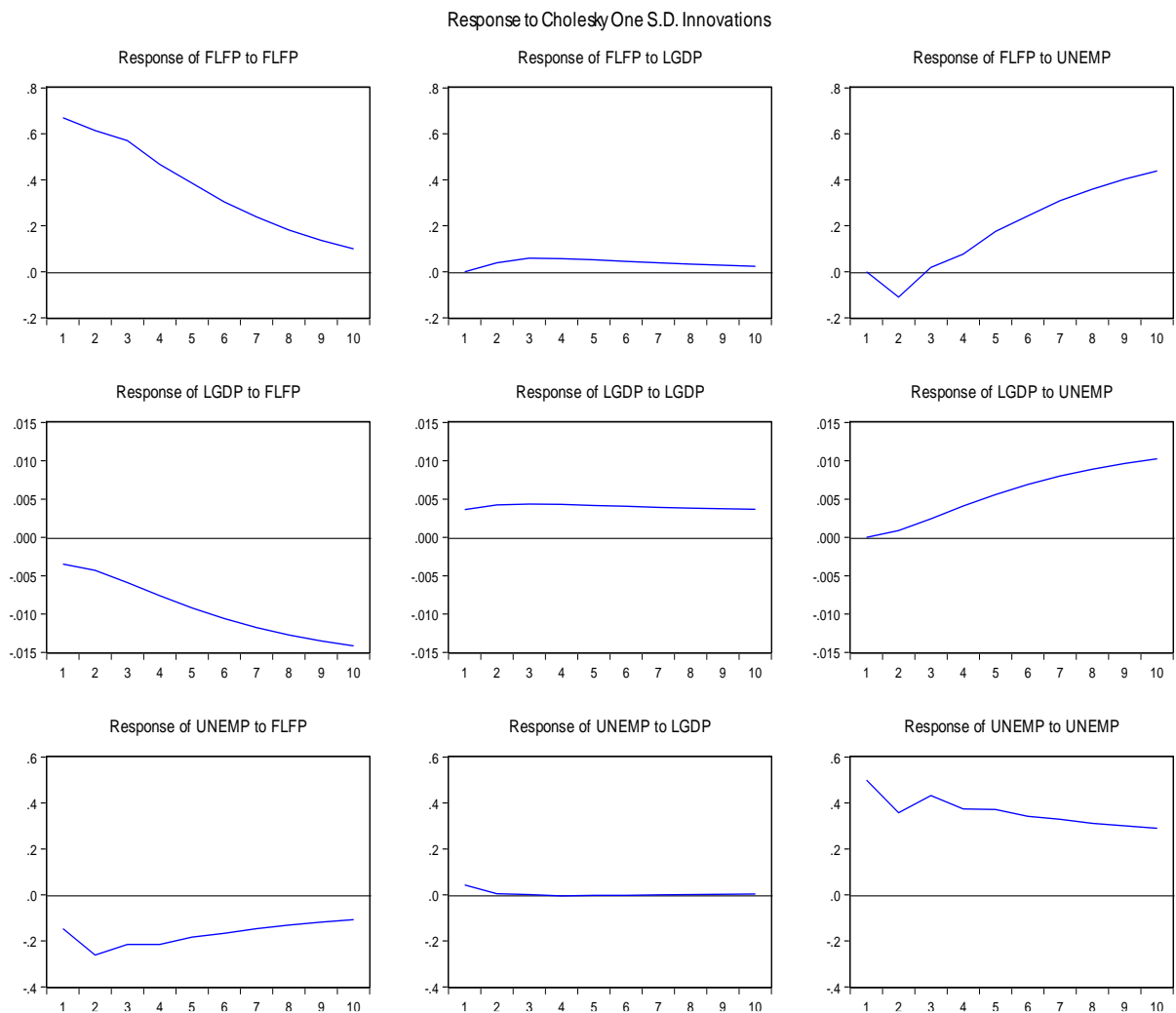
According to the above equation, the short-run coefficient (-0.008) is insignificant but with a correct sign, see Appendix A (Table 8). It is actually indicating that the behaviour of UNEMP will not assist FLFP to return to its equilibrium level in the short run. The short-run coefficient of the equation indicates that only FLFP has a significant negative impact on UNEMP at the 10% level of significance. Given the coefficient of -0.26, a 1% rise in FLFP will result in a 0.26 decrease in UNEMP.

The log of GDP is statistically insignificant with a negative sign indicating that it might have a negative effect on UNEMP. It is also realistic in the economy of South Africa which is labour-intensive. In other words, as the country's economic growth increases, unemployment falls.

#### **5.5.4.4. VECM Impulse Response**

The following figures are the results of the IRF analysis of the behaviour of each variable in the VECM model as per the hypothesis of the study.

**Figure 5.3: VECM Impulse Response**



The first row indicates the response of FLFP to a one standard deviation shock on itself, LGDP and UNEMP. The first diagram actually shows that FLFP declines for every one standard deviation shock on itself. When analysing the effect of a shock of LGDP on FLFP, FLFP responds positively up to the third horizon, and declines thereafter to close to the zero line. Moreover, it indicates the  $\cap$ -shape relationship between the two variables. With one standard deviation shock on UNEMP, FLFP responded negatively throughout to the second horizon after it became positive at an increasing rate.

The second row indicates the response of LGDP to a shock on FLFP, a shock on itself and UNEMP. The log of GDP responded negatively to a one unit shock in FLFP. It is declining throughout the horizon. Furthermore, a one unit standard deviation on UNEMP results in an increase in LGDP throughout the horizon. On the other hand,

the last row shows the results of the response of UNEMP to a one unit shock on FLFP, LGDP and UNEMP itself. UNEMP responds negatively to a one unit shock on FLFP, and to a shock in LGDP. It is negative to the second quarter and becomes zero throughout the horizon.

#### **5.5.4.5. VECM diagnostic**

VECM diagnostic is presented for serial correlation, normality and heteroscedasticity. The results are shown in Appendix A (Table A8). The LM serial correlation test with the null hypothesis of no serial correlation is accepted at the selected lag length. The probability values are far greater than 5%; in other words, we fail to reject the null hypothesis. White heteroscedasticity (only levels and squares) test has been adopted in testing heteroscedasticity among residuals given the null hypothesis that residuals are heteroscedastic. By observing the probability values of the test which is greater than 5%, the null hypothesis is rejected in favor of the alternative which state that residuals are homoscedastic. Normality tests (Lutkepohl) are applied with a null hypothesis stating that residuals are normally distributed. In this case, the results (Table A8, iii) reveal that, jointly, the residuals are normally distributed. The null hypothesis fails to be rejected.

#### **5.5.4.6. Granger Causality**

Multiple relationships in the direction of causality among variables may exist. Therefore, a Granger Causality test is conducted to assist in indicating the direction of causality in the exogenous and endogenous variables in the model. Granger (1988), Guechari (2012) and Ogutu (2014) indicate that Granger Causality is concerned with the short-run forecastability while cointegration is more concerned about the long-run equilibrium relationship among variables. The causality test is tested under the null hypothesis of non-causality between variables. The decision of causality is taken based on the significant of the probability values of the test. The study rejects the null hypothesis if the probability value is less than 10%.

Pair-wise Granger Causality is presented in Table 5.6 below which highlights short-run results using AIC information criteria to select lag length. Lag length in these results is 2.

<b>Table 5.6: Granger Causality Test</b>			
Null Hypothesis:	Obs	F-Statistic	Prob.
LGDP does not Granger Cause FLFP	33	6.76631	0.0040
FLFP does not Granger Cause LGDP		3.05714	0.0630
UNEMP does not Granger Cause FLFP	33	3.90162	0.0320
FLFP does not Granger Cause UNEMP		2.63709	0.0893
UNEMP does not Granger Cause LGDP	33	7.86230	0.0020
LGDP does not Granger Cause UNEMP		0.50754	0.6074

Source: *Eviews 9. Student version*

In the table, F-statistics and probabilities of respective variables are used to check whether we reject the null hypothesis of no causality among variables in consideration or not, is reported. When studying the causality relationship of FLFP and the log of GDP, we reject the null hypothesis and conclude that LGDP does cause FLFP (first hypothesis). Importantly, when studying the reverse hypothesis that FLFP does not Granger cause LGDP is also rejected infavor of an alternative indicating that FLFP does Granger cause LGDP. Hence, there is a bi-directional effect between FLFP and log GDP in the short-run which is basically shown by respective probability values that enable us to reject the null hypothesis of no causality in the first hypotheses. Therefore, hypothesis 3 fails to be rejected.

At 10% level of significance, both null hypotheses are rejected when studying the direction of causality between FLFP and UNEMP in the South African context. Given the probability values of less that 10%, a bi-directional Granger causality is observed. Unemployment and FLFP do cause each other in the short run. UNEMP appears to Granger cause LGDP. However, the hypothesis that LGDP does not Granger cause UNEMP is accepted in this case. Thus, there is a uni-directional causality relationship running from unemployment to LGDP.

## 5.6. SINGLE EQUATIONS ESTIMATIONS

This section is concerned with the estimation of the FM-OLS, CCR and DOLS using yearly data collected from the database sources outlines in Chapter 4. It observes the

nature of the long and short-run relationship between FLFP, economic growth and unemployment in South Africa.

**Table 5.6: Descriptive statistics for yearly data (1980–2015)**

	<b>Mean</b>	<b>Standard-Dev</b>	<b>Minimum</b>	<b>Maximum</b>	<b>Obs.</b>	<b>Countries</b>
<b>FLFP</b>	20.815	5.121	9.241	27.800	35	1
<b>LGDP</b>	10.782	0.089	10.654	10.939	35	1
<b>UNEMP</b>	39.884	3.663	33.973	44.090	35	1
	<b>Skewness</b>	<b>Kurtosis</b>	<b>Jarque-Ber</b>	<b>Probability</b>	<b>Obs.</b>	<b>Countries</b>
<b>FLFP</b>	-0.800	2.603	4.084	0.129	36	1
<b>LGDP</b>	0.424	1.913	2.852	0.240	36	1
<b>UNEMP</b>	-0.394	1.466	4.461	0.107	36	1

Source: Eviews

The summary of measures of centrality and dispersion, kurtosis and skewness are provided in the above table. Measures of centrality are therefore computed with variables after logarithmic transformations. FLFP and UNEMP indicate that they are skewed to the left, as shown by a negative skewness coefficient. The FLFP and UNEMP kurtosis values are 4.08 and 4.46. A normal distribution is shown by LGDP with a coefficient of 2.85 relatively closer to 3 which indicate a perfectly normal distribution. The JB statistic which measures normality of variables is reported in this table with respect to its probability values under the null hypothesis that variables are normally distributed. According to the results, we accept the null hypothesis and conclude that all variables are normally distributed, the probability values are greater than 0.05%. However, this provides evidence that outliers are not a problem in each individual samples of the data set. These results are in line with the quarterly time-series descriptive statistics results.

Apart from descriptive statistics, stationarity of all variables in the dataset is computed. The summary (Appendix B, Table B1) of these tests with the ADF uses Schwarz Information Criterion for optimal lag lengths selection while the bandwidths for the PP-tests are indicated using the Newey-West Bartlett Kernel selection as indicated above in the stationarity of quarterly time-series. The asterisks \*, \*\* and \*\*\* indicate statistical significance at 1%, 5% and 10% respectively. The results indicates that the series is non-stationary at levels for all variables. after first differencing all of them to become stationary, thus variables are integrated of order one, I(1).

### 5.6.1. Fully-Modified Ordinary Least Squares (FMOLS)

Table B1 (Appendix B) presents the long-run relationship of the estimated FMOLS regression, the cointegration equation is as follows:

$$FLFP_t = 97.83 - 6.11LGDP_t + 0.23UNEMP_t + 4.29D1995 + 1.74D2002$$

(5.4)

The results reveal that there is a positive significant long-run relationship between FLFP and UNEMP and the shocks that took place in 1995 and 2002. The results indicate that a 1% rise in economic growth lowers the participation of women in the labour force by 6.1%. In other words, low economic growth induces high women participation in the labour force. This finding supports the hypothesis mentioned in Chapter 2. It is noticeable that with an increase in a country's output, women lower their participation up to a certain level whereafter it begins to rise again (Forgha & Mbella, 2016). Holding all other variables constant, a 1% rise in UNEMP leads to a 0.23 increase in FLFP. This is in line with the theory discussed in Chapter 2.

Holding GDP and UNEMP constant, a shock that took place after democracy raised the level of FLFP in the labour force about to 4.3%.

**Table 5.7: Fully-Modified Ordinary Least Squares (FMOLS)**

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LGDP	-6.109745	2.784956	-2.193839	0.0361
UNEMP	0.230915	0.046666	4.948275	0.0000
C	97.83240	30.32513	3.226116	0.0030
D1995	4.296636	0.411333	10.44564	0.0000
D2002	1.740831	0.618458	2.814791	0.0085
R-squared	0.968846	Mean dependent var		40.05347
Adjusted R-squared	0.964692	S.D. dependent var		3.571496
S.E. of regression	0.671099	Sum squared resid		13.51120
Long-run variance	0.528369			

#### 5.6.1.1. FMOLS Diagnostic Checks

The model diagnostic is done to ensure its reliability. The variation in the participation of women in the labour force is explained fully by natural log of GDP, UNEMP and the 1995 & 2002 shocks with 0.97% ( $R^2$  and adjusted  $R^2$ ). Figure C and Table B2

(Appendix B) show the results of the normality and cointegration tests indicating that residuals are normally distributed and there is long-run relationship among variables.

### 5.6.2. Canonical Cointegration Regression (CCR)

The long-run equilibrium equation is presented below obtained from the CCR estimation and confirms the results obtained from the FM-OLS estimate (Table B1, Appendix B). It indicates that explanatory variables are statistically significant in driving the participation of women in the labour force in South Africa in the long run, hence there is cointegration among the variables.

$$FLFP_t = 106.34 - 6.89LGDP_t + 0.22UNEMP_t + 4.33D1995 + 1.9D2002 \quad (5.5)$$

The above regression results show that at an early level of economic growth increasing investment in economic infrastructure by 1% will lead to a 0.09% increase in economic growth at a 1% level of significance, everything else being constant. In contrast, a 1% rise in the level of unemployment causes female participation in the labour market to fall by 0.28%. Hence, the results confirm those reported by the FM-OLS which indicate a decrease of 0.23%. The signs of the coefficients are consistent with economic theory, except that of social infrastructure investment, which suggests that it takes away from growth in the long run. Additionally, all coefficients are statistically significant except UNEMP at conventional significance levels.

**Table 5.8: Canonical Cointegrating Regression (CCR)**

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LGDP	-6.890676	3.224513	-2.136966	0.0409
UNEMP	0.222278	0.052798	4.209985	0.0002
C	106.3410	35.17506	3.023193	0.0051
D1995	4.337934	0.464230	9.344356	0.0000
D2002	1.904871	0.718649	2.650626	0.0127
R-squared	0.968355	Mean dependent var		40.05347
Adjusted R-squared	0.964135	S.D. dependent var		3.571496
S.E. of regression	0.676371	Sum squared resid		13.72433
Long-run variance	0.528369			

### 5.6.2.1. CCR diagnostic tests results

The overall model indicates 96% goodness of fit, meaning that about 96% of the variations in FLFP are explained by variations in economic growth and unemployment respectively. Figure B2 and Tables B4 (Appendix B) shows that the CCR model's residuals are normally distributed and both cointegration tests fail to reject the null hypothesis of no cointegration at 5% level of significant. Moreover, cointegration tests confirm the long-run relationship among variables (see Table B5).

### 5.6.3. Dynamic Ordinary Least Squares

DOLS results in Table B5 below indicate that female's participation is derived by the high level of unemployment in South Africa, their participation in the labour force responds positively and significantly during the high rate of unemployment and also to the shock that took place in 1995. The long-run equilibrium equation is given by:

$$FLFP_t = 36.16 - 0.49LGDP_t + 0.33UNEMP_t + 3.D1995 + 0.29D2002 \quad (5.6)$$

The above equation (5.8) shows that a 1% increase in rate of unemployment, *ceteris paribus*, would actually result in a 0.33% rise in the participation of women in the labour force. The long-run coefficients for private investment, social infrastructure investment and employment of 0.03, 0.28 and -0.07 are all statistically insignificant; thus, these variables do not have a long-run relationship with growth. These results therefore confirm those attained from both the FM-OLS and CCR approaches with regard to the statistically significant correlation between economic infrastructure investments and economic growth.

**Table 5.9: Dynamic Ordinary Least Squares (DOLS)**

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LGDP	-0.497543	3.529548	-0.140965	0.8892
UNEMP	0.333485	0.064716	5.153080	0.0000
C	36.16051	38.37742	0.942234	0.3563
D1995	3.403780	0.605501	5.621431	0.0000
D2002	0.288974	0.799928	0.361250	0.7214
R-squared	0.979414	Mean dependent var		40.15987
Adjusted R-squared	0.970057	S.D. dependent var		3.499410
S.E. of regression	0.605544	Sum squared resid		8.067031
Long-run variance	0.475084			

### 5.6.3.1. DOLS Diagnostic Checking

Diagnostic checking of the model is performed. There is a 98% goodness of fit, indicating that about 96% of the variations in FLFP is explained by variations in economic growth and unemployment respectively. Tables B6 (Appendix B) shows that residuals in the DOLS model are normally distributed and both cointegration tests fail to reject the null hypothesis of no cointegration at 5% level of significance. Hence, there is a long-run relationship among variables.

### 5.7. GRANGER CAUSALITY TEST (1980 – 2015 TIME-SERIES)

The Granger Causality results below were analysed using AIC information criteria to select lag length of 2. When studying the causality relationship of FLFP and the log of GDP, we reject the null hypothesis in favour of the alternative hypothesis and conclude that LGDP causes FLFP and FLFP causes LGDP.

<b>Table 5.10: Pair-wise Granger Causality Tests</b>			
Date: 01/14/19 Time: 22:00			
Sample: 1980 2015			
Lags: 2			
Null Hypothesis:	Obs	F-Statistic	Prob.
LGDP does not Granger Cause UNEMP	34	0.08396	0.9197
UNEMP does not Granger Cause LGDP		5.97797	0.0067
FLFP does not Granger Cause UNEMP	34	1.21585	0.3111
UNEMP does not Granger Cause FLFP		1.35873	0.2729
FLFP does not Granger Cause LGDP	34	6.81083	0.0038
LGDP does not Granger Cause FLFP		4.02724	0.0286

Source: Eviews

## 5.8. CHAPTER SUMMARY

This chapter mainly concentrated on the core part of the study in estimating all regression models as proposed in Chapter 4, and interpretations with respect to every model estimated were provided in this chapter. The analysis began with descriptive statistics and unit-root testing. Under descriptive analysis, graphical plots of all variables in their levels and first differences, measures of centrality and dispersion, and covariance-correlation matrix were computed. The unit-root testing included ADF and PP-tests presented in a table format and briefly interpreted. The regression model, cointegration and error correction model are conducted and clearly interpreted. The last section provided diagnostic inspections and model stability. Therefore, in accordance with the value chain of this study, all steps have been followed and fully interpreted although the long-run Granger Causality is not accounted for as it requires the application of the VAR model of which is beyond the scope of the project.

## **CHAPTER 6: SUMMARY, POLICY RECOMMENDATIONS AND LIMITATIONS**

### **6.1. INTRODUCTION**

This concluding chapter provides a summary of the study as well as the main findings and its implications for policy, as well as the limitations of the study and future directions that similar research could take to overcome these limitations.

This chapter summarises the paper whose aim was to investigate the long-run and short-run relationship between FLFP, LGDP and UNEMP. First, the study employed quarterly data ranging from 1980 Q1 to 2016 Q3 within the VAR/VECM framework to model the long and short-run relationships between the variables. Second, the study used yearly data from 2008 to 2015, within the single equation cointegrating framework to confirm the existence of long-run cointegrating relationships between the variables. The rest of this chapter consists of three sections. Section 6.1 provides a summary and a reflection on the main findings, section 6.2 discusses the policy recommendations based on the findings, while section 6.3 addresses the limitations of the study and makes recommendations for future studies.

### **6.2. SUMMARY OF THE STUDY**

The first chapter provided the objectives and hypothesis of the study. Moreover, it discussed the problem statement and briefly discussed the main objective of each chapter. As noted above, the aim of this study was to investigate the long and short-run relationships that existed between FLFP, the UNEMP and LGDP in South Africa. For this purpose, Chapters 2 and 3 discussed theories and empirical literature which underpin this relationship. Theories are categorised into two parts, that is, labour supply theories and conceptual framework. Labour supply theories covered in this study included the theory of the decision to work, the lifecycle model and HCT. These theories explained in general the supply and demand for labour but were limited in explaining specifically the participation of women in the labour force; hence, Chapter 2 extended the conceptual framework to include discussions from the empirical literature conducted in different countries by various researchers using different methodologies on both micro and macro data. Some studies conducted in South Africa reported findings that women were deprived of participating in the labour market due to their role as a caregiver, i.e., being a housewife or raising children and thus they depended on their husband's income and only participated in labour on family farms

and in low-paying jobs. International studies confirm the U-shape relationship between FLFP and economic growth. They indicate that during low levels of economic development, more women enter the labour market, but as the economy develops, they leave the market for better education and rejoin when the country is well developed. Related to the U-theory is that when economic growth is low, unemployment will be high and under such circumstances women will be more inclined to enter the workforce. This study included LGDP and UNEMP to determine whether these theoretical constructs could be tested empirically.

Chapter 4 discussed the methodological framework adopted to examine the long and short-run relationships between FLFP, LGDP and UNEMP. It began with a description of the sources and units of measurement of the data and included an outline of the descriptive statistics used to assess the characteristics of the data. It was followed by a full discussion of spurious regression and non-stationarity concepts. The unit-root concept was segregated into three subsections that discussed relevant tests such as the DF, ADF and PP tests. This was followed by discussing the cointegration concept and VAR/VECM modeling approach and how it is utilised to estimate short run as well as long-run cointegrating relationships. In regard to the short-run VAR relationships, the chapter outlined the Impulse Response Function System (IRFS) and variance decomposition techniques used to analyse the relationships between the variables. Moreover, the chapter explained the single equation methods used to analyse the long-run cointegration relationships involving the yearly data.

Chapter 5 presented the findings of the study. **With regard to quarterly data, through the observation of the VAR-generated IRFs, it was found that in** the short-run FLFP is nonresponsive to an unanticipated shock to LGDP. This result is corroborated by the long-run cointegrating regression results where the coefficient associated with LGDP was found to be statistically insignificant. These results were supported by the variance decomposition results. This result was supported by the DOLS findings which estimated the long-run cointegrating relationship between the three variables using yearly data. However, these findings were contradicted by the findings of the FM-OLS and CCR single equation results which employed yearly data and found that FLFP responds negatively to LGDP, which is consistent with the prediction of the U theory. The study is inclined to accord a greater weighting to the FM-OLS and CCR results

since they were based on a larger dataset (1980-2015) while the VAR/VECM results were based on a much smaller data set, viz., 2008-2016 quarterly data.

Furthermore, the quarterly data VAR model found that, in the short-run, FLFP is highly responsive in the positive direction to an unanticipated shock to the unemployment rate after 1.5 quarters and keeps rising into the future. This result is supported by the long-run analysis, the Johansen long-run cointegration findings as well as the three single equation results that suggest a statistically significant cointegration relationship between the variables. The single equation analysis suggests the elasticity lies between 0.22-0.33% while the Johansen long-run elasticity is 2.7%. This study is inclined to support the single equation results due to the size of the dataset as noted above. Hence the U theory implication governing the relationship between female participation in the workforce and the unemployment rate is supported.

### **6.3. POLICY IMPLICATIONS**

Since the study has confirmed a negative relationship between FLFP and economic growth, it implies that South African society possesses a traditionalist mindset that the role of women is primarily that of a home executive, which reduces their vocational contributions to society. Hence, policy makers should initiate projects and provide incentives that encourage women to participate in the workforce on an equal footing with men in order to achieve their full potential.

The study also confirmed that when unemployment is rising, more women tend to participate in the labour force. This is probably because of the need to supplement the family income; however, they enter the workforce at a lower wage rate. Part of the unemployment problem in South Africa is high wage rates in the formal economy. South Africa must follow a flexible wage rate policy which will attract more unskilled and semi-skilled individuals into the workforce.

### **6.4. LIMITATIONS AND RECOMMENDATIONS OF THE STUDY**

First, the study employed a small quarterly dataset to assess a VAR/VECM model which requires a more substantial dataset that covers at least a full business cycle. Future studies ought to accumulate a much larger dataset and rerun the analysis. If only small datasets are available, then future studies ought to pursue a panel

VAR/VECM (as well as panel single equation methods) analysis through involving more countries in the study.

Second, the study limited itself to merely pursuing a cointegration analysis between FLFP, LGDP and unemployment rate. This is a serious limitation as the full implications of the U theory were not investigated. Specifically, the study did not investigate the threshold effects of GDP, i.e., the study did not incorporate LDGP<sup>2</sup> into the model to test for turning points in levels of GDP growth. Future studies ought to address this omission. Similarly, threshold effects for unemployment must also be tested in future studies.

Third, the study investigated cointegration between only three variables. It did not include an important variable like the female literacy rate which is likely to have a significant impact on the participation of women in the workforce. This could be a topic for a future study.

## REFERENCES

- Abdullah, N. & Bakar, N. A. (2011). The causal relationship between fertility and women labor force participation: evidence for the four selected ASEAN countries. *European Journal of Social Sciences*, 26 (2): 154-158.
- Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the institute of Statistical Mathematics*, 21 (1): 243-247.
- Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the institute of Statistical Mathematics*, 21(1), 243-247.
- Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, 60(2): 255-265.
- Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, 60(2), 255-265.
- Akaike, H. (1979). A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika*, 66 (2): 237-242.
- Akaike, H. (1979). A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika*, 66(2), 237-242.
- Alam, S., (2010). A reassessment of the effects of exchange rate volatility on Pakistan's export demand: ARDL approach. *European Journal of Economics, Finance and Administrative Sciences*, 21.
- Amador, D., Bernal, R. & Peña, X. (2013). The rise in female participation in Colombia: Fertility, marital status or education? *Ensayos sobre Política Económica*, 31(71): 54-63.
- Asgar, Z. & Abid, I. (2007). *Performance of lag length selection criteria in three different situations*. Available at: <https://mpra.ub.uni-muenchen.de/40042/>
- Asgar, Z., & Abid, I. (2007). Performance of lag length selection criteria in three different situations. Available at: <https://mpra.ub.uni-muenchen.de/40042/>
- Asteriou, D. and Hall, S. (2007) *Applied Econometrics A Modern Approach*. Palgrave Macmillan, New York
- Ateino, R. (2006). *Government policy and female labour force participation in Kenya: Implications for poverty reduction*. Nairobi: Institute for Development Studies, University of Nairobi.
- Banerjee, O., Ghaoui, L.E. and d'Aspremont, A.,(2008). Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data. *Journal of Machine learning research*, 9(Mar), pp.485-516.

- Barker, F. (2013). *The South African labour market. Critical issues for renaissance*. Pretoria: Van Schaik.
- Bhalla, S. and Kaur, R. (2011) Labour force participation of women in India: some facts, some queries. Working Paper, 40.
- Bhalla, S., & Kaur, R. (2011). Labour force participation of women in India: some facts, some queries. *Asia Research Centre Working Paper 40*. Available at: <http://eprints.lse.ac.uk/38367/1/ARCWP40-BhallaKaur.pdf>
- Blau, F. D. & Kahn, L. M. (2013). *Female labor supply: Why is the US falling behind?* (No. w18702). Cambridge: National Bureau of Economic Research.
- Brooks, C. (2014). *Introductory econometrics for finance*. London: Cambridge University.
- Brown, R.L, et al.(1975). Techniques for Testing the Constancy of Regression Relationships over Time : *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 37, No. 2
- Cameron, N., Preece, M.A. and Cole, T.J.,(2005). Catch-up growth or regression to the mean? Recovery from stunting revisited. *American Journal of Human Biology: The Official Journal of the Human Biology Association*, 17(4), pp.412-417
- Casale, D. (2004). What has the feminization of the labour market 'bought' women in South Africa? Trends in labour force participation, employment and earnings, 1995–2001. *Journal of Interdisciplinary Economics*, 15 (3-4), 251-275.
- Casale, D. M. (2003). *The rise in female labour force participation in South Africa: an analysis of household survey data, 1995-2001* Unpublished doctoral dissertation. University of KwaZulu-Natal, Durban.
- Chamlou, N. Muzi, S. & Ahmed, H. (2011). Understanding the determinants of female labor force participation in the Middle East and North Africa region: The role of education and social norms in Amman. *Working Papers 31*. Bologna: AlmaLaurea Inter-University Consortium.
- Chen, J., Shao, X., Murtaza, G. & Zhao, Z. (2014). Factors that influence female labor force supply in China. *Economic Modelling*, 37, 485-491.
- Cheung, Y.W and Lai K S. (1993) A Fractional Cointegration Analysis of Purchasing Power Parity. *Journal of Business & Economic Statistics*, vol. 11, issue 1, 103-12

- Collins, S. M., Bosworth, B.P. & Rodrik, D. (1996). Economic growth in East Asia: accumulation versus assimilation. *Brookings Papers on Economic Activity*, 1996 (2): 135-203.
- Collins, S.M., & Bosworth. B., (1996). Economic Growth in East Asia: Accumulation versus Assimilation. *Brookings Papers on Economic Activity.*, vol. 27, issue 2, 135-204
- Contreras, D., Puentes, E. & Bravo, D. (2005). Female labour force participation in greater Santiago, Chile: 1957–1997. A synthetic cohort analysis. *Journal of International Development*, 17 (2): 169-186.
- Das, M. S., Jain-Chandra, S., Kochhar, M.K. & Kumar, N. (2015). *Women workers in India: why so few among so many?(No. 15-55)*. Washington, D.C.: International Monetary Fund.
- De-Graft Acquah, H. (2010). Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship. *Journal of Development and Agricultural Economics*, 2(1): 001-006.
- Dickey, D.A. and Fuller, W.A. (1979) Distribution of the Estimates for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*
- Dickey, D.A. and Fuller, W.A. (1981) Likelihood Ratio Statistics for Autoregressive Time Series with Unit Root. *Econometrica*, 49, 1057-1072.
- Ejaz, M. (2007). Determinants of female labor force participation in Pakistan: An empirical analysis of PSLM (2004-05) micro data. *The Lahore Journal of Economics*, 12: 203-235.
- Engle, R.F. and Granger, C.W.J. (1987) Cointegration and Error Correction Representation, Estimation and Testing. *Econometrica*, 55, 251-276.
- Evans, M.D.R. & Kelly, J. (2008). Trends in woman's labour force participation in Australia. *Social Science Research*, 37 (8): 287-310.
- Even, W.E., & Macpherson, D.A. (1990). The gender gap in pensions and wages. *The Review of Economics and Statistics*, 72 (2): 259-265.

- Fatima, A. & Sultana, H. (2008). Tracing out the U-shape relationship between female labour force participation rate and economic development for Pakistan. *International Journal of Social Economics*. 36 (1/2): 182-198.
- Festus, L., Kasongo, A., Moses, M. & Yu, D. (2016). The South African labour market, 1995–2015. *Development Southern Africa*, 33 (5): 579-599.
- Forgha, N.G. & Mbella, M.E. (2016). The implication of female labour force participation on economic growth in Cameroon. *International Journal of Development and Economic Sustainability. European Centre for Research Training and Development, United Kingdom*, 4 (1): 34-47.
- Gaddis, I. & Klasen, S. (2014). Economic development, structural change, and women's labor force participation. *Journal of Population Economics*, 27 (3): 639-681.
- Garavan, T.N., Morley, M., Gunnigle, P. & Collins, E. (2001). Human capital accumulation: the role of human resource development. *Journal of European Industrial Training*, 25 (2/3/4): 48-68.
- Goksel, I. (2013). Female labour force participation in Turkey: The role of conservatism. *Woman's Studies International Forum*, 41(13): 45-54.
- Goldin, C (1990): *Understanding the Gender Gap: An Economic History of American Women* (US: Oxford University Press).
- Goldin, C. (1995). The U-shaped female labor force function in economic development in economic history. In Schultz, T.P. (Ed.). *Investment in Women's Human Capital*. Chicago: University of Chicago Press. 61-90.
- Granger, C.W. (1988). Some recent development in a concept of causality. *Journal of Econometrics*, 39 (1-2): 199-211.
- Granger, C.W., 1981. Some properties of time series data and their use in econometric model specification. *Journal of econometrics*, 16(1), pp.121-130.
- Guechari, Y. (2012). An empirical study on the effects of real effective exchange rate on Algeria's trade balance. *International Journal of Financial Research*, 3 (4): 102.

- Gujarati, D.N. & Porter, D.C. (2009). *Basic econometrics*. International Edition. Singapore: McGraw-Hill.
- Gujarati, D.N. and Porter, D.C. (2009) *Basic Econometrics*. 5th Edition, McGraw-Hill/Irwin, Boston, 400.
- Hannan, E. J. & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*. 190-195.
- Hannan, E. J. & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 190-195.
- Harris, R. and Sollis, R. (2003) *Applied Time Series Modelling and Forecasting*. Wiley: Chichester. ISBN 9780470844434
- Hudu, I. (2014). *Assessing female labour force participation rate in the construction industry* Unpublished doctoral dissertation. Kwame Nkrumah University of Science and Technology, Ashanti.
- Jaumotte, F. (2003). *Female labour force participation: Past trends and main determinants in OECD countries*. *OECD Economics Department Working Papers, No. 376*. Paris: OECD Publishing.
- Johnston, J. & DiNardo, J. (1972). *Econometric methods (Vol. 2)*. New York: McGraw-Hill.
- Johnston, J., & DiNardo, J. (1972). *Econometric methods (Vol. 2)*. New York: McGraw-Hill.
- Killingsworth, M.R. (1983). *Labor supply*. New York: Cambridge University Press.
- Kinoshita, Y. & Guo, F. (2015). *What can boost female labor force participation in Asia?* Washington D.C.: International Monetary Fund.
- Klasen, S. & Pieters, J. (2015). *What explains the stagnation of female labor force participation in urban India?* Available at: <https://openknowledge.worldbank.org/bitstream/handle/10986/21668/WPS7222.pdf?sequence=2>

- Lechman, E. & Kaur, H. (2015). Economic growth and female labor force participation –verifying the U-feminization hypothesis. New evidence for 162 countries over the period 1990-2012. *Economics and Sociology*, 8 (1): 246-257.
- Lechman, E. (2014). Female labour force participation and economic growth – Re-examination of U-shaped curve. *GUT FME Working Paper Series A, No. 3/2014(21)*. Gdańsk: Faculty of Management and Economics, Gdańsk University of Technology.
- Liew, V. K. S. (2004). Which lag length selection criteria should we employ? *Economics Bulletin*, 3 (33): 1–9
- Lincove, J.A. (2008). Growth, girls' education, and female labor: A longitudinal analysis. *The Journal of Developing Areas*, 41 (2): 45-68.
- MacKinnon, J.G., (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of applied econometrics*, 11(6), pp.601-618.
- MacKinnon, J.G., Haug, A.A. and Michelis, L. (1999) Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration. *Journal of Applied Econometrics*
- Makepe, P.M. & Oageng, P.T. (2012). Determinants of female labour force participation in agricultural sector in Botswana. *Asian-African Journal of Economics and Econometrics*, 12 (2): 449-461.
- Marimuthu, M., Arokiasamy, L. & Ismail, M. (2009). Human capital development and its impact on firm performance: Evidence from developmental economics. *Journal of international Social Research*, 2(8).
- Mazalliu, M. & Zogjani, J., 2015. The effect of female labour force in economic growth and sustainability in transition economies - case study for SEE countries. *Academic Journal of Business, Administration, Law and Social Science*, 1 (2): 179-186.
- Mishra, V. & Smyth, R (2009). Female Labour force participation and total fertility rate in OECD: New evidence from panel Cointegration and Granger Casualty testing. *Journal of Economics & Business*, 62 (2010): 48-64.

- Morrill, M.S. & Morrill, T. (2012). Intergenerational links in female labour force participation. *Labour Economics*, 20 (2013): 38-47
- Moses, E.,(2016). Quality of education and the labour market: A conceptual and literature overview. Stellenbosh Economic Paper:07/11
- Mujahid, N. & uz Zafar, N. (2012). Economic growth-female labour force participation nexus: an empirical evidence for Pakistan. *The Pakistan Development Review*, 51 (4): 565-585.
- Mullis, I.V., Martin, M.O., Foy,P. & Arora, A.,(2012). *TIMSS2011 international results in mathematics*. International Association for the Evaluation of Educational Achievement. Herengracht 487, Amsterdam,1017 BT, the Netherlands.
- Mutsau, I. (2012). Empirical analysis of money demand in South Africa. Unpublished doctoral thesis. Department of Economics. University of Zululand, Empangeni.
- Mutsau, I., (2013). Empirical analysis of money demand in South Africa (1980-2011): an autoregressive distributed lag approach (Doctoral dissertation, University of Zululand).
- Nafziger, E. W. (2012). *Economic development*. (5<sup>th</sup> ed.). London: Cambridge University Press.
- Ntuli, M. & Wittenberg, M. (2013). Determinants of black women's labour force participation in post-apartheid South Africa. *Journal of African Economies*, 22 (3): 347-374.
- Ntuli, M. (2007). *Determinants of South African women's labour force participation, 1995-2004*. Available at: <http://ftp.iza.org/dp3119.pdf>
- Olivetti, C. (2013). The female labour force and long-run development: The American experience in comparative perspective. In *Human Capital in History: The American Record*. Cambridge: National Bureau of Economic Research. 161-197
- P. Takaendesa and N.M. Odhiambo. (2007) *Studies in Economics and Econometrics*, Volume 31, Issue 3,

- Phillips, P.C., 1987. Towards a unified asymptotic theory for autoregression. *Biometrika*, 74(3), pp.535-547.
- Phillips, R. & Perron, P. (1988) Testing for a unit root time series regression. *Biometrika*, 75 (2): 335-346.
- Pieters, J. & Klasen, S. (2011) Drivers of female labour force participation in urban India during Indian's economic boom. *Proceedings of the Germany Development Economics Conference, Berlin. no.65*. Available at: [https://www.econstor.eu/bitstream/10419/48322/1/65\\_pieters.pdf](https://www.econstor.eu/bitstream/10419/48322/1/65_pieters.pdf)
- Posel, D. & Casele, D. (2002). Gender aggregate: Women subsistence farmers effect the unemployment count. *Agenda*, 49: 82-88.

## REFERENCE

- S. Das, S. Jain-Chandra, K. Kochhar, N. Kumar.(2015). Women workers in India: why so few among so many. IMF Working Paper. , 1-31.
- Sackey, H.A. (2005). Female labour force participation in Ghana, the effects of education. *AERC research paper 150*, Nairobi: The Regal Press.
- Schwarz, G., (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), pp.461-464.
- Siphambe, H. & Motswapong, M. (2010). Female participation in the labour market of Botswana: Results from the 2005/06 labour force survey data. *Botswana Journal of Economics*, 7 (11): 65-78.
- Sorsa, P., Mares, J., Didier, M., Guimaraes, C., Rabate, M., Tang, G., & Tuske, A. (2015). Determinants of the low female labour force participation in India. *OECD Economics Department Working Papers, No. forthcoming*. Paris: OECD Publishing.
- Stock J, Watson MW.(1993). A Simple Estimator of Cointegrating Vectors in Higher-Order Integrated Systems. *Econometrica*. Vol.61 No.4 :783-820
- Sudarshan, RM. And Bhattacharya, S. (2000)Through the Magnifying Glass: Women's Work and Labour Force Participation in Urban . vol XLIV no 48.

- Tam, H. (2011). U-shaped female labor participation with economic development: Some panel data evidence. *Economics Letters*, 110 (2), 140-142.
- Tandrayen-Ragoobur, V., Ummersingh, S. & Bundhoo, Y. (2011). The power to choose: women and labour market decisions in Mauritius. *Journal of Emerging Trends in Economics and Management Sciences*, 2(3), 193-205.
- Thévenon, O. (2013). Drivers of female labour force participation in the OECD. *OECD Social, Employment and Migration Working Papers, No. 145*. Paris: OECD Publishing. <http://dx.doi.org/10.1787/5k46cvrgnms6-en>
- Tsani, S., Paroussos, L., Fragiadakis, C., Charalambidis, I. & Capros, P. (2012). *Female labour force participation and economic development in Southern Mediterranean countries: What scenarios for 2030?* Brussels: MedPro.
- van der Westhuizen, C., Goga, S., Ncube, L., & Oosthuizen, Morné. (2007). Women in the South African Labour Market, 1995 - 2005. SSRN Electronic Journal. 10.2139/ssrn.967277.
- Varian, H.R., (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28 (2): 3-28.
- Verick, S. (2014). *Female labor force participation in developing countries*. Bonn: IZA World of Labor.
- Warunsiri, S. & McNown, R. (2010). Female labor supply in Thailand: 1985-2004. A synthetic cohort analysis. *IBS Working Paper*. Boulder: Colorado
- Wilkinson, F. (Ed.). (2013). *The dynamics of labour market segmentation.*: Elsevier.
- Yakubu, Y.A., 2010. Factors influencing female labor force participation in South Africa in 2008. *The African Statistical Journal*, 11.
- Ziramba, E. (2007) Demand for money and expenditure components in South Africa: assessment from unrestricted error-correction models. *South African Journal of Economics*, 75(3),412-424.

## APPENDIX B: SINGLE EQUATIONS RESULTS

### APPENDIX A: MULTIVARIATE EQUATIONS FIGURES AND TABLES

The following multivariate equations figures were derived from having analysed the quarterly data: 2008 Q3 -2016 Q3.

Figure A1 (i): Stationary series: 2008Q3-2016Q3

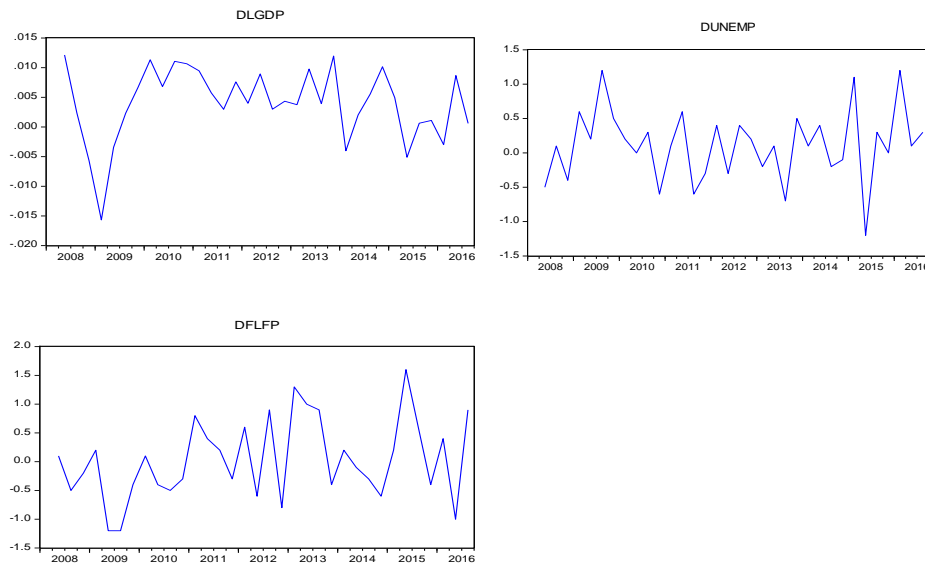


TABLE A1: Optimal Lag Length Results (2008Q3-2016Q3)

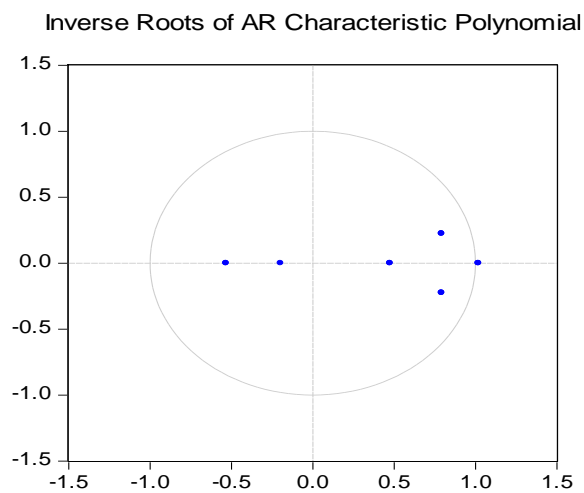
Lag	LogL	LR	FPE	AIC	SC	HQ
0	-26.31913	NA	0.001331	1.891557	2.030330	1.936793
1	82.27990	189.1725*	2.16e-06*	-4.534187	-3.979095*	-4.353241*
2	91.40566	14.13021	2.19e-06	-4.542300*	-3.570890	-4.225645
3	97.09189	7.703932	2.84e-06	-4.328509	-2.940780	-3.876144
4	106.6014	11.04335	3.02e-06	-4.361384	-2.557335	-3.773309

**Table A2: Autoregressive Roots Results (2008Q3-2016Q3)**

Root	Modulus
1.018256	1.018256
0.793337 - 0.224644i	0.824529
0.793337 + 0.224644i	0.824529
-0.531494	0.531494
0.475318	0.475318
-0.197157	0.197157

Warning: At least one root outside the unit circle.  
VAR does not satisfy the stability condition.

**Figure A2: Autoregressive Roots Results**



**Table A3: VAR Estimation Results (2008Q3-2016Q3)**

Vector Autoregression Estimates

Date: 02/07/17 Time: 01:07

Sample (adjusted): 2008Q3 2016Q3

Included observations: 33 after adjustments

Standard errors in ( ) &amp; t-statistics in [ ]

	FLFP	LGDP	UNEMP
FLFP(-1)	0.704529 (0.21111) [ 3.33724]	0.001352 (0.00165) [ 0.81979]	-0.282807 (0.16580) [-1.70573]
FLFP(-2)	0.068269 (0.17628) [ 0.38727]	-0.001504 (0.00138) [-1.09189]	0.321593 (0.13845) [ 2.32288]
LGDP(-1)	28.92411 (22.2593) [ 1.29941]	1.048341 (0.17393) [ 6.02721]	-21.54188 (17.4816) [-1.23226]
LGDP(-2)	-16.65289 (22.3183) [-0.74615]	-0.135765 (0.17440) [-0.77849]	21.47265 (17.5279) [ 1.22506]
UNEMP(-1)	-0.282373 (0.22617) [-1.24852]	0.002183 (0.00177) [ 1.23539]	0.598728 (0.17762) [ 3.37080]
UNEMP(-2)	0.101031 (0.26868) [ 0.37603]	0.003263 (0.00210) [ 1.55422]	0.389612 (0.21101) [ 1.84641]
C	-166.7784 (72.2274) [-2.30907]	1.176959 (0.56439) [ 2.08538]	-0.345247 (56.7245) [-0.00609]
R-squared	0.893760	0.992340	0.809450
Adj. R-squared	0.869243	0.990572	0.765478
Sum sq. resids	9.948545	0.000607	6.136157
S.E. equation	0.618576	0.004834	0.485805
F-statistic	36.45466	561.3869	18.40791
Log-likelihood	-27.04013	133.0705	-19.06687
Akaike AIC	2.063038	-7.640634	1.579811

Schwarz SC	2.380479	-7.323193	1.897252
Mean dependent	49.65455	14.87594	24.78485
S.D. dependent	1.710645	0.049782	1.003158

**Table A4: VAR Diagnostic (2008Q3-2016Q3)**

i).VAR Residual Serial Correlation

LMTests

Sample:2008Q3 2016Q3

Included observations: 33

Lags	LM-Stat	Prob
1	7.073784	0.6294
2	2.963856	0.9657
3	8.832337	0.4529

ii)VAR Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)

Joint test:

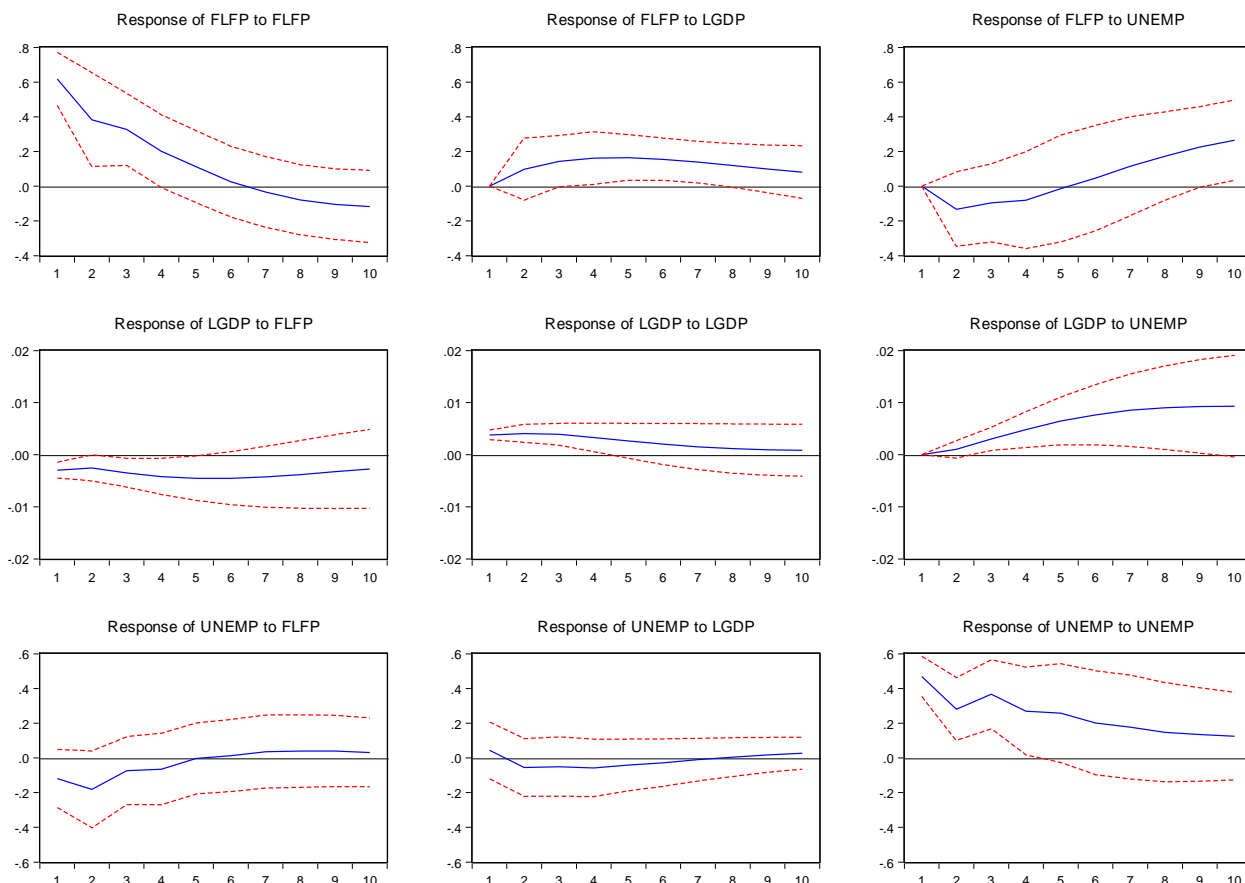
Chi-sq	df	Prob.
71.15922	72	0.5059

**iii). Normality Test**

Component	Jarque-Bera	Df	Prob.
1	0.644789	2	0.7244
2	0.119007	2	0.9422
3	0.210164	2	0.9003
Joint	0.973960	6	0.9866

**Figure A3: VAR Impulse Response Function (2008Q3-2016Q3)**

Response to Cholesky One S.D. Innovations  $\pm 2$  S.E.



**Table A5: Variance Decomposition of FLFP (2008Q3-2016Q3)**

Period	S.E.	FLFP	LGDP	UNEMP
1	0.618576	100.0000	0.000000	0.000000
2	0.745912	95.13526	1.712754	3.151984
3	0.832593	91.80237	4.349654	3.847977
4	0.875413	88.32736	7.343735	4.328903
5	0.897965	85.49620	10.36760	4.136191
6	0.912737	82.83297	12.90524	4.261795
7	0.930956	79.75649	14.61832	5.625188
8	0.957621	76.05044	15.35183	8.597733
9	0.994259	71.65133	15.23794	13.11073
10	1.038920	66.89522	14.57195	18.53283

**Table A6: Variance Decomposition of LGDP**

Period	S.E.	FLFP	LGDP	UNEMP
--------	------	------	------	-------

1	0.004834	38.21945	61.78055	0.000000
2	0.006897	32.50312	65.29311	2.203761
3	0.009173	32.69933	55.08933	12.21134
4	0.011656	33.11550	42.14256	24.74194
5	0.014321	31.86822	31.32348	36.80830
6	0.016988	29.76536	23.67882	46.55582
7	0.019539	27.24346	18.51365	54.24288
8	0.021889	24.73393	15.03753	60.22854
9	0.024012	22.40395	12.65150	64.94455
10	0.025911	20.34163	10.97137	68.68700

**Table A7: Variance Decomposition of UNEMP**

Period	S.E.	FLFP	LGDP	UNEMP
1	0.485805	6.005630	0.800000	93.19437
2	0.592482	13.45788	1.425678	85.11644
3	0.702238	10.68271	1.529259	87.78803
4	0.757204	9.912954	1.898282	88.18876
5	0.800799	8.864963	1.958251	89.17679
6	0.826297	8.352316	1.952530	89.69515
7	0.845880	8.150537	1.877250	89.97221
8	0.859610	8.098639	1.820188	90.08117
9	0.871070	8.090801	1.811686	90.09751
10	0.880853	8.040937	1.864012	90.09505

**Table A8: Cointegration Test**

Unrestricted Cointegration Rank Test (Trace)

Hypothesized		Trace	0.05	
No. of CE(s)	Eigenvalue	Statistic	Critical Value	Prob.**
None *	0.567297	44.84444	35.19275	0.0034
At most 1	0.257865	17.20019	20.26184	0.1252
At most 2	0.199881	7.358810	9.164546	0.1087

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level

\* denotes rejection of the hypothesis at the 0.05 level

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

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized		Max-Eigen	0.05	
No. of CE(s)	Eigenvalue	Statistic	Critical Value	Prob.**
None *	0.567297	27.64426	22.29962	0.0082
At most 1	0.257865	9.841379	15.89210	0.3490
At most 2	0.199881	7.358810	9.164546	0.1087

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level

\* denotes rejection of the hypothesis at the 0.05 level

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

**Table A9: VECM (2008Q3-2016Q3)**

Standard errors in ( ) &amp; t-statistics in [ ]

Cointegrating Eq:	CointEq1
FLFP(-1)	1.000000
LGDP(-1)	6.278565 (13.3424) [ 0.47057]
UNEMP(-1)	-2.716367 (0.72371) [-3.75337]

Standard errors in ( ) & t-statistics in [ ]

Cointegrating Eq:		CointEq1		
C		-78.16722		
		(185.570)		
		[-0.42123]		
Error Correction:		D(FLFP)	D(LGDP)	D(UNEMP)
CointEq1		-0.058127	-0.001215	-0.007936
		(0.04651)	(0.00035)	(0.03625)
		[-1.24979]	[-3.49323]	[-0.21894]
D(FLFP(-1))		-0.005036	0.001100	-0.263663
		(0.18142)	(0.00136)	(0.14139)
		[-0.02776]	[ 0.81033]	[-1.86483]
D(LGDP(-1))		13.76835	0.152721	-7.214651
		(22.4848)	(0.16819)	(17.5235)
		[ 0.61234]	[ 0.90803]	[-0.41171]
D(UNEMP(-1))		-0.378363	-0.001519	-0.305974
		(0.25098)	(0.00188)	(0.19560)
		[-1.50754]	[-0.80917]	[-1.56427]
R-squared		0.194406	0.373883	0.094419
Adj. R-squared		0.111069	0.309113	0.000738
Sum sq. resids		13.04525	0.000730	7.923558
S.E. equation		0.670698	0.005017	0.522710
F-statistic		2.332759	5.772413	1.007882
Log-likelihood		-31.51160	130.0402	-23.28496
Akaike AIC		2.152218	-7.638798	1.653634
Schwarz SC		2.333613	-7.457404	1.835029
Mean dependent		0.033333	0.003735	0.130303
S.D. dependent		0.711366	0.006036	0.522903

**Table A10: OLS Estimate**

System: UNTITLED  
 Estimation Method: Least Squares  
 Date: 03/24/17 Time: 22:36  
 Sample: 2008Q3 2016Q3  
 Included observations: 33  
 Total system (balanced) observations 99

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.058127	0.046509	-1.249786	0.2147
C(2)	-0.005036	0.181416	-0.027757	0.9779
C(3)	13.76835	22.48476	0.612342	0.5419
C(4)	-0.378363	0.250980	-1.507544	0.1353
C(5)	-0.001215	0.000348	-3.493234	0.0008
C(6)	0.001100	0.001357	0.810328	0.4200
C(7)	0.152721	0.168188	0.908035	0.3664
C(8)	-0.001519	0.001877	-0.809175	0.4206
C(9)	-0.007936	0.036247	-0.218939	0.8272
C(10)	-0.263663	0.141387	-1.864830	0.0656
C(11)	-7.214651	17.52355	-0.411712	0.6816
C(12)	-0.305974	0.195601	-1.564274	0.1214
Determinant residual covariance		1.00E-06		

## VECM Diagnostic

Table A11: VEC Residual Serial Correlation LM Tests

Lags	LM-Stat	Prob
1	11.15136	0.2655
2	6.643619	0.6742
3	9.357522	0.4049

**Table A12: VEC Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)**

Joint test:		
Chi-sq	df	Prob.
74.36700	84	0.7648

**Table A13: VECM Normality**

Component	Jarque-Bera	Df	Prob.
1	1.004565	2	0.6051
2	0.074526	2	0.9634
3	0.412166	2	0.8138
Joint	1.491256	6	0.9601

Dependent Variable: D(FLFP)

Method: Least Squares (Gauss-Newton / Marquardt steps)

Date: 03/27/17 Time: 13:37

Sample (adjusted): 2008Q3 2016Q3

Included observations: 33 after adjustments

$D(FLFP) = C(1) * (FLFP(-1) + 6.27856525822 * LGDP(-1) - 2.71636653724 * UNEMP(-1) - 78.1672170514) + C(2) * D(FLFP(-1)) + C(3) * D(LGDP(-1)) + C(4) * D(UNEMP(-1))$

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.058127	0.046509	-1.249786	0.2214
C(2)	-0.005036	0.181416	-0.027757	0.9780
C(3)	13.76835	22.48476	0.612342	0.5451
C(4)	-0.378363	0.250980	-1.507544	0.1425
R-squared	0.194406	Mean dependent var		0.033333
Adjusted R-squared	0.111069	S.D. dependent var		0.711366
S.E. of regression	0.670698	Akaike info criterion		2.152218
Sum squared resid	13.04525	Schwarz criterion		2.333613
Log likelihood	-31.51160	Hannan-Quinn criter.		2.213252
Durbin-Watson stat	1.932335			

Wald Test:  
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	1.219328	(3, 29)	0.3204
Chi-square	3.657985	3	0.3008

Null Hypothesis: C(2)=C(3)=C(4)=0  
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	-0.005036	0.181416
C(3)	13.76835	22.48476
C(4)	-0.378363	0.250980

Dependent Variable: D(LGDP)

Method: Least Squares (Gauss-Newton / Marquardt steps)

Date: 03/27/17 Time: 13:20

Sample (adjusted): 2008Q3 2016Q3

Included observations: 33 after adjustments

D(LGDP) = C(5)\*( FLFP(-1) + 6.27856525822\*LGDP(-1) - 2.71636653724  
\*UNEMP(-1) - 78.1672170514 ) + C(6)\*D(FLFP(-1)) + C(7)\*D(LGDP(-1)) + C(8)\*D(UNEMP(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
C(5)	-0.001215	0.000348	-3.493234	0.0016
C(6)	0.001100	0.001357	0.810328	0.4244
C(7)	0.152721	0.168188	0.908035	0.3713
C(8)	-0.001519	0.001877	-0.809175	0.4250

R-squared	0.373883	Mean dependent var	0.003735
Adjusted R-squared	0.309113	S.D. dependent var	0.006036
S.E. of regression	0.005017	Akaike info criterion	-7.638798
Sum squared resid	0.000730	Schwarz criterion	-7.457404
Log likelihood	130.0402	Hannan-Quinn criter.	-7.577765
Durbin-Watson stat	2.015241		

Dependent Variable: D(UNEMP)

Method: Least Squares (Gauss-Newton / Marquardt steps)

Date: 03/27/17 Time: 13:35

Sample (adjusted): 2008Q3 2016Q3

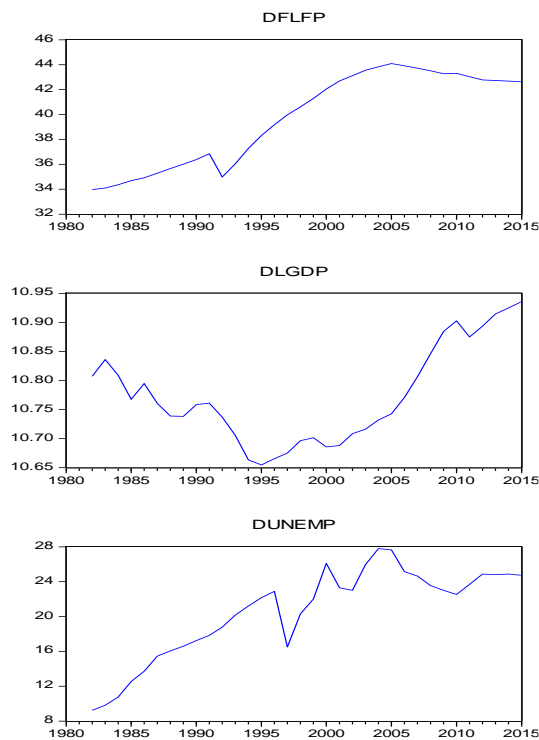
Included observations: 33 after adjustments

$D(UNEMP) = C(9) * (FLFP(-1) + 6.27856525822 * LGDP(-1) - 2.71636653724 * UNEMP(-1) - 78.1672170514) + C(10) * D(FLFP(-1)) + C(11) * D(LGDP(-1)) + C(12) * D(UNEMP(-1))$

	Coefficient	Std. Error	t-Statistic	Prob.
C(9)	-0.007936	0.036247	-0.218939	0.8282
C(10)	-0.263663	0.141387	-1.864830	0.0724
C(11)	-7.214651	17.52355	-0.411712	0.6836
C(12)	-0.305974	0.195601	-1.564274	0.1286
R-squared	0.094419	Mean dependent var		0.130303
Adjusted R-squared	0.000738	S.D. dependent var		0.522903
S.E. of regression	0.522710	Akaike info criterion		1.653634
Sum squared resid	7.923558	Schwarz criterion		1.835029
Log likelihood	-23.28496	Hannan-Quinn criter.		1.714668
Durbin-Watson stat	2.024974			

The following tables and graphs were compiled through the analysis of the yearly data: 1980-2015.

**Figure A1(ii): Stationary series: 1980-2015**



**Table B1: Summary of Unit-Roots Results**

Variable	Model	ADF test		Critical Values		PP test		Critical Values							
		La g	$\tau_\tau, \tau_\mu, \tau$	1%	5%	10%	BW	3	0.4971	5%	10%				
FLFP	Intercept	0	0.4203	1%	-3.6329	5%	-2.9484	10%	-2.6128	3	0.4971	5%	-2.9484	10%	-2.6128
	Intercept & trend	0	0.9944	1%	-4.2436	5%	-3.5443	10%	-3.2047	3	0.9839	5%	-3.5443	10%	-3.2047
	None	0	0.9955	1%	-2.6327	5%	-1.9507	10%	-1.6111	3	0.9816	5%	-1.9507	10%	-1.6111
	Intercept	0	0.0020	1%	-3.6390	5%	-2.9511	3	0.0017	5%	-2.9511				

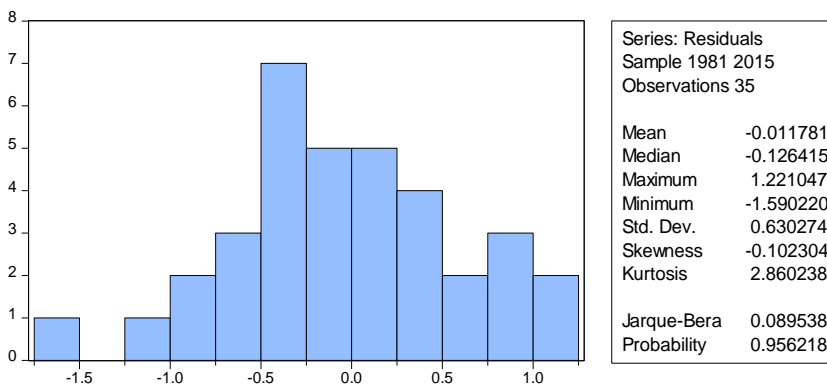
Variable	Model	ADF test		Critical Values		PP test		Critical Values	
Dflfp				10%	-2.6134			10%	-2.6134
	Intercept & trend	0	0.0038	1% 5% 10%	-4.2528 -3.5484 -3.2070	2	0.0038	1% 5% 10%	-4.2528 -3.5484 -3.2070
	None	0	0.0004	1% 5% 10%	-2.6347 -1.9510 -1.6109	3	0.0004	1% 5% 10%	-2.6347 -1.9510 -1.6109
<b>Conclusion: non-stationary, integrated of order one I(1)</b>									
LGDP	Intercept	1	0.8762	1% 5% 10%	-3.6394 -2.9511 -2.6143	3	0.9306	1% 5% 10%	-3.6329 -2.9484 -2.6128
	Intercept & trend	1	0.5025	1% 5% 10%	-4.2528 -3.5484 -3.2070	3	0.8862	1% 5% 10%	-4.2436 -3.5442 -3.2047
	None	1	0.8006	1% 5% 10%	-2.6347 -1.9510 -1.6109	3	0.8697	1% 5% 10%	-2.6326 -1.9507 -1.6110
DIgdp	Intercept	0	0.0043	1% 5% 10%	-3.6394 -2.9511 -2.6143	2	0.0037	1% 5% 10%	-3.6394 -2.9511 -2.6143
	Intercept & trend	0	0.0021	1% 5% 10%	-4.2528 -3.5484 -3.2070	3	0.0021	1% 5% 10%	-4.2528 -3.5484 -3.2070
	None	0	0.0002	1% 5% 10%	-2.6347 -1.9510 -1.6109	2	0.0002	1% 5% 10%	-2.6347 -1.9510 -1.6109
<b>Conclusion: non-stationary, integrated of order one I(1)</b>									
UNEMP	Intercept	0	0.2030	1% 5% 10%	-3.6329 -2.9484 -2.6128	20	0.0375	1% 5% 10%	-3.6329 -2.9484 -2.6128
	Intercept & trend	0	0.4074	1% 5% 10%	-4.2436 -3.5442 -3.2046	23	0.6568	1% 5% 10%	-4.2436 -3.5442 -3.2046
	None	0	0.9019	1% 5% 10%	-2.6326 -1.9506 -1.6110	30	0.9441	1% 5% 10%	-2.6326 -1.9506 -1.6110
	Intercept		0.0000	1%	-3.6394			1%	-3.6394

Variable	Model	ADF test		Critical Values			PP test		Critical Values	
				5%	10%		5%	10%	5%	10%
Dunemp		0		5%	-2.9511	33	0.0000	5%	-2.9511	
				10%	-2.6143			10%	-2.6143	
	Intercept & trend	0	0.0003	1%	-4.2732	19	0.0000	1%	-4.2732	
				5%	-3.5577			5%	-3.5577	
				10%	-3.2123			10%	-3.2123	
	None	0	0.0000	1%	-2.6327	10	0.0000	1%	-2.6347	
				5%	-1.9510			5%	-1.9510	
				10%	-1.6109			10%	-1.6109	

Conclusion: non-stationary, integrated of order one I(1)

Source: Researcher's estimation.

**Table B2: FM-OLS Normality Test**



**Table B3: FM-OLS Cointegration Tests**

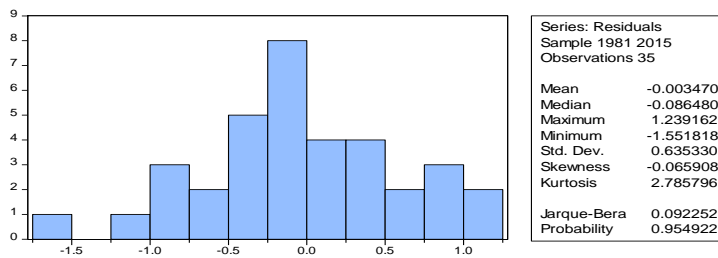
I) Engle-Granger

	Value	Prob.*
Engle-Granger tau-statistic	-3.750813	0.0803
Engle-Granger z-statistic	-20.80833	0.0599

II) Phillips-Ouliaris

	Value	Prob.*
Phillips-Ouliaris tau-statistic	-3.705383	0.0874
Phillips-Ouliaris z-statistic	-18.98403	0.0954

**Table B4: CCR Normality Test**



**Table B5: CCR Cointegration Tests**

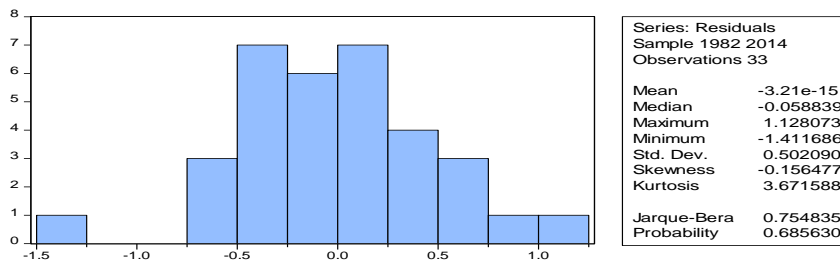
I) Engle-Granger

	Value	Prob.*
Engle-Granger tau-statistic	-3.750813	0.0803
Engle-Granger z-statistic	-20.80833	0.0599

II) Phillips-Ouliaris

	Value	Prob.*
Phillips-Ouliaris tau-statistic	-3.705383	0.0874
Phillips-Ouliaris z-statistic	-18.98403	0.0954

**Table B6: DOLS Normality Test**



**Table B7: DOLS Cointegration Tests**

I) Engle-Granger

	Value	Prob.*
Engle-Granger tau-statistic	-3.750813	0.0803
Engle-Granger z-statistic	-20.80833	0.0599

II) Phillips-Outliaris

	Value	Prob.*
Phillips-Ouliaris tau-statistic	-3.705383	0.0874
Phillips-Ouliaris z-statistic	-18.98403	0.0954