An Empirical Analysis of Energy Demand in Sub-Saharan Africa

By

Aisha Kolawole

A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy of Oxford Brookes University

Oxford Brookes Business School

Oxford Brookes University, UK

July 2017

ACKNOWLEDGEMENTS

I would like to thank and express deep gratitude to my supervisors, Dr Sola Adesola and Professor Glauco De Vita, who guided and helped me with their expertise to complete this PhD study. They provided me with academic insights, commitment and care throughout this research. Without their expert supervision and intellectual inputs, this thesis would not have been possible. I am also grateful for their constructive criticism and interest in my work. I shall always remain indebted to them for their encouragement, confidence and kindness towards me. My sincere gratitude also goes to Dr David Bowen for his understanding and support during my research.

I am grateful to Oxford Brookes University for giving me bursary during my study, and my caring parents who also made extra resources available. Without this support, I would have been unable to embark on the PhD journey. Also, I would like to appreciate my very supportive husband and best friend, Damilola Ekundayo for always encouraging me to work hard on my research. Lastly, my appreciation goes to my loving son, Ikeoluwa Ajayi for his understanding and support throughout my study period and always.

ABSTRACT

The thesis presents the first comprehensive analysis of energy demand in Sub-Saharan Africa (SSA) by analysing the demand functions for both aggregated and disaggregated (by energy types) energy demand models. The main aim of the study is to explore the impact of income, price, economic structure, urbanisation and population on the demand for energy in SSA for the period 1980 to 2014. To achieve this aim, the study adopts a panel data model approach to analyse secondary data sourced from publicly available and widely used energy and economic databases.

The aggregate demand model analysis reveals that income, urbanisation and energy prices are significant drivers of aggregate energy demand in the long run in SSA. The panel model was analysed using the panel cointegration technique. From the results, there is evidence that as consumers earn more, they are able to acquire more energy gadgets and appliances. Similarly, with economic growth, both new and existing firms can expand their production scale which increases the overall amount of energy consumed. The results also suggest that a rise in rural-urban migration increase the total energy consumed, as consumers move towards the use of modern energy equipment that is more accessible in urban areas. The results also indicate that, in accordance with economic theory and the law of demand, an increase in the price of energy reduces the total amount of energy consumed, though such response is found to be fairly inelastic.

For the disaggregated models, the specific individual energy types analysed are: electricity, petrol, diesel, liquefied petroleum gas (LPG), kerosene and solid biomass. From the panel linear static models employed, the analysis shows that economic structure, urbanisation, population and income are all significant drivers of the demand for the energy types analysed in SSA. This suggests that as population increases, there will be an increase in demand for each energy type in the region. The same response applies to an increase in urban population and income. From the results, this study found that population is the predominant factor behind the increase in demand for the analysed energy types, with the highest elasticity. The results are in line with the theory of demand.

The identified factors, their analysed impacts on the demand for energy and the reported elasticities, whilst increasing our academic knowledge of the main determinants of energy in SSA, can also help policymakers prepare evidence-based and more effective energy demand management, to meet the energy need of consumers in the region. The findings suggest the need for stringent energy conservation policies through effective energy efficiency practice in all the sixteen countries analysed, to ensure that an increase in energy use does not lead to more greenhouse gas (GHG) emission and the produced energy is well utilized. Furthermore, there is a need for increased competition through the use of independent power companies to improve energy service delivery and markets in Sub-Saharan Africa.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

AfDB	African Development Bank
EIA	Energy Information Administration
GDP	Gross Domestic Product
GGFR	Global Gas Flaring Reduction Partnership
IEA	International Energy Agency
IPCC	Intergovernmental Panel on Climate Change
LAPSSET	Lamu Port Southern Sudan-Ethiopia Transport
LNG	Liquefied Natural Gas
LPG	Liquefied Petroleum Gas
NEPAD	New Partnership for Africa's Development
OPEC	Organisation of the Petroleum Exporting Countries
SAPP	Southern Africa Power Pool
SASOL	South African Synthetic Oil Limited
SSA	Sub-Saharan Africa
UNFPA	United Nations Population Fund
WAGPCO	West African Gas Pipeline Company
WDI	World Development Indicators
WNA	World Nuclear Association

Chapter 1 : Introduction

1.1 Chapter overview

This chapter starts with a brief overview of the context and background of the issue investigated in this PhD thesis. The research aim and objectives to answer the main research question are clearly stated, alongside the contribution to knowledge. The chapter concludes by introducing the overall structure of the thesis.

1.2 Context and background

It has been well articulated by academics that energy demand modeling plays a crucial role in effective energy planning, strategy formulation and sound energy policy recommendations (Bhattacharyya and Timilsina, 2010). The theory of demand provides a useful account of how changes in income and other factors, which influence the demand for energy, can be modelled. This has been used to model energy demand in both developed and developing countries.

Energy is an integral part of a modern economy, because it is an important factor of production employed in the production of many goods and services, when combined with capital and labour (Keho, 2016). It is also the main factor behind global warming due to greenhouse gas emissions. The projected increase from 46% to 58% in global energy demand is expected to be predominantly from developing countries (EIA, 2007). This may explain why the past decade has seen a rapid development of energy demand models in many developing countries, including Sub-Saharan Africa (SSA). The increase in energy demand modelling in SSA is due to many other compelling reasons.

SSA population accounts for 14% of the global population but only 4% of the total energy consumed (IEA, 2014). Most of the energy consumed is derived from solid biomass like fuelwood and charcoal which accounts for 75% of the total energy consumed in the region (Lambe *et al.,* 2015). The consumption of solid biomass (see Figure 1.1) used by the residential sector for cooking has adverse effects on the environment and health (Rahut *et al.*, 2016). Despite the adverse health implications, wood fuel use for cooking and heating is still on the increase across the region (Sulaiman *et al.,* 2017). The associated health issue includes the risk of severe respiratory infections in the under 5 age group, lung cancer, and bad pregnancy outcomes. At both the micro and macro level, there are adverse implications linked to the use of solid biomass for cooking. At the macro level, research has shown that in 2010 for example, US$12 billion was lost in SSA due to the health, environmental and economic costs by households to biomass use (Lambe *et al.,* 2015). Furthermore, at the micro and household level, the time used by women and young girls for collecting solid fuel could have been used for other productive activities. The time spent varies across the region, with an average of 2.1 hours per day, while the time spent in Sierra Leone is as high as 5 hours per day (Rysankova *et al.,* 2014). Even in urban areas of SSA, solid biomass is still widely used in the form of charcoal.

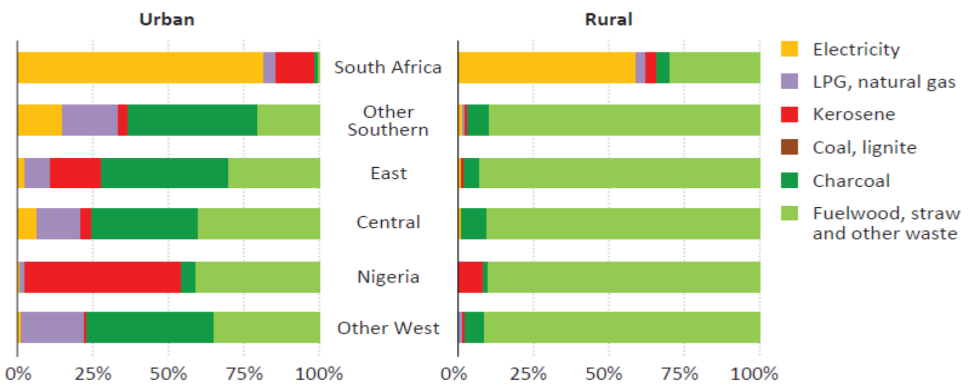

Figure 1.1: Household sources of cooking fuel
(**Source:** Rysankova *et al.,* 2014)

The majority of the population of 1.01 billion (World Bank, 2015) in SSA live in rural areas. However, the last few decades have recorded an increase in urban population, from 22.1% in 1980 to 38% in 2015 (World Bank, 2015). The slump in commodity prices has also reduced the rate of growth of GDP (the percentage annual change of the size of the economy in terms of the dollar amount of goods and services produced) from an average of 6% recorded in the last decade to 3.5%

in 2015 (Newiak, 2016). According to Newiak (2016), a growth rate of 4% is projected for 2017, due to low commodity prices (World Bank, 2016).

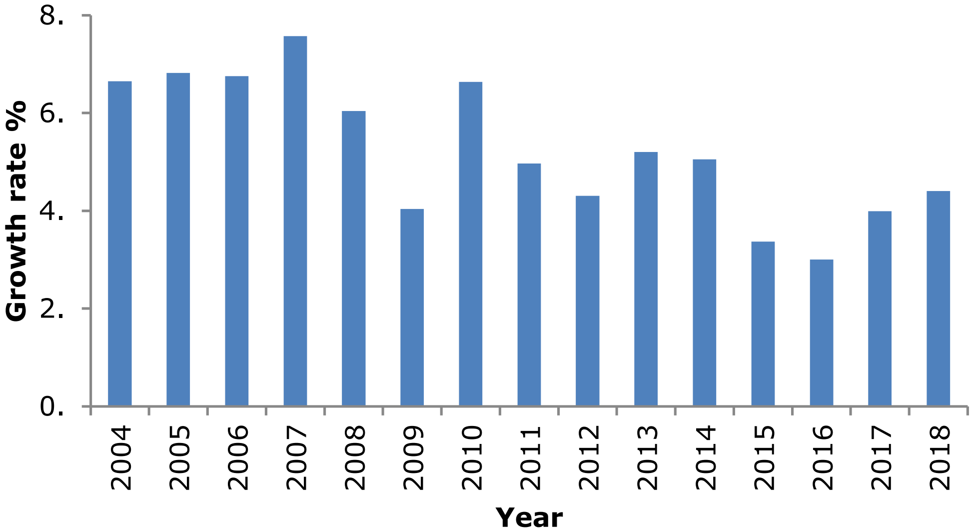

Figure 1.2: Real GDP growth rate in SSA

(**Data Source:** Newiak, 2016)

The reported economic growth rates (see Figure 1.2) have not led to an increase in per capita income or the standard of living. The region has remained the poorest region in the world due to a higher increase in population than GDP growth (Newiak, 2016). The top 10 countries with the highest fertility rates in the world are in SSA (PRB, 2013). The number of young people between the age of 10 and 24 in SSA is the highest in the world, currently at 324.5 million, a number expected to increase to 436 million by 2025 (UNFPA, 2012). An increase in energy supply would reduce the economic hardship in the region because it would provide more opportunities for the youthful population. Modern energy can lead to the creation of more jobs (also through self employment) and good industrial and manufacturing sectors.

Of course, energy is a 'necessary but not sufficient' condition for the economic growth of a country (Bildirici, 2013). Access to modern energy is needed to foster economic growth, social development and healthy living (IEA, 2015). Studies have predicted that there will be a $15 increase in the Sub-Sahara African GDP for each dollar invested in the power sector (IEA, 2014). The power sector in SSA is the least developed when compared to other regions of the world, with a national electrification rate of only 32% (IEA, 2013).

Evidence has also shown that people shift towards more modern energy sources as their income increases (see, for example, Brew-Hammond, 2007). An adequate power supply is a critical condition for sustainable development in the region, which is still afflicted by widespread poverty (70% of people in Sub-Saharan Africa live on less than US$2 per day, UNICEF 2014), high unemployment and a deficient healthcare system. Struggling small and medium scale enterprises (SMEs) and the few industries in the region would also perform better through a more adequate power supply, which would, in turn, aid the social and economic development of the region. In other words, for sustainable development in the region, energy is needed (Kaygusuz, 2012). This has been the focus of some research in developing countries, including Sub-Saharan African countries with several authors exploring the causality between economic growth and energy consumption (see, for example, the studies reviewed in Section 3.8 of Chapter 3, including Kraft and Kraft, 1978; Mensah, 2014; Akinlo, 2009; Odhiambo, 2009).

1.3 The research gap

A single empirical study that analyses both the cross-country aggregate energy demand and disaggregate demand by energy type in SSA, will provide information on the main determinants of energy demand. The findings of the study will guide academics, governments, policy makers and investors on how to meet the energy need of consumers so as to improve the standard of living, and also enhance SSA socio-economic development. Although a number of mostly single-country studies have investigated the determinants of energy demand in some African countries, no single study has examined the elasticities of energy demand at aggregate level and by energy type using the latest panel model techniques, for a large proposition of representative- SSA countries.

In addition, the present PhD study incorporates two important variables that have not been considered in previous studies that investigated energy demand in the region, namely, economic structure and the degree of urbanisation; both of which are important when considering the current trends and dynamics in SSA (as the results of this study will demonstrate). Furthermore, this is the most up-to-date study with a sample period ending in 2014 for the aggregate energy demand study in Sub-Saharan Africa. Likewise, it is the most up-to-date study for disaggregated energy demand by energy types in SSA, with a sample ending in 2013.

1.4 Research aim and objectives

The primary aim of this doctoral research is to fill the above gap in the literature by identifying and analysing the factors driving energy demand in SSA at aggregate and disaggregated level, by energy type. The empirical analysis investigates the impact of income, price, urbanisation, economic structure and population on aggregate and disaggregated (fuel type) energy demand in SSA. The specific objectives of the research are:

(i) To provide a full and comprehensive analysis of the energy sources in the region;
(ii) To critically review both the theoretical and empirical literature on energy demand (energy consumption) in developing regions including SSA;
(iii) To provide a review of the aggregate and disaggregated pattern of energy demand literature in the region;
(iv) To provide preliminary conclusions on the main energy issues and prospects affecting the region;
(v) To develop a comprehensive econometric model for the analysis of a cross-country aggregate and disaggregated (energy type) energy demand function for SSA to estimate elasticities of energy demand to changes in the explanatory variables;

(vi) To make an original contribution to the existing body of knowledge on energy demand in SSA and draw out relevant policy implications in light of the research findings.

The focus of the research is Sub-Saharan Africa (SSA), which comprises of 47 countries but 16 are analysed due to data constraints. The issue investigated is one which deserves attention so as to realise the economic and developmental potential of the region, and build the infrastructure needed to end energy poverty. Also, considering the urgent need for most of the countries to diversify from total dependence on mineral export, providing sufficient energy would open up more opportunities. Guided by a positivist approach, theory-based, *a priori* hypotheses are used and tested to verify and establish the demand for energy in SSA in this study.

1.5 Contribution to knowledge

This research makes an original contribution to knowledge because, to date, there is no single study of up-to date aggregate and disaggregated energy demand in Sub-Saharan Africa (SSA). The study analyses the demand for petrol, diesel, liquefied petroleum gas (LPG), electricity, kerosene and solid biomass in SSA. In other words, the impact of income, energy prices, urbanisation, population and economic structure on the demand for aggregate and disaggregate energy in SSA are investigated and presented in this PhD thesis. Secondary data from the International Energy Agency, the World Bank and the International Monetary Fund databases, which span over 33 years for 16 countries in the region, are used for the panel data econometric analysis.

The findings can aid the development of an appropriate policy framework for meeting the energy need of consumers in the region while also enabling informed investment decisions in the development of interregional capital intensive energy systems across SSA. Especially when considering cash constraints and the expected future growth of energy consumption in developing countries, it is reasonable for policy makers and planners to be interested in evidence that can guide decision making. It follows that information on present energy consumption patterns and the significant determinants of energy demand are critical tools for both investment and planning decisions in SSA. Therefore, the policy recommendations and implications that stem from the findings of the results of the study will guide governments across the region on how to reduce energy poverty across the region, and to meet the energy need of consumers.

1.6 Thesis structure

The structure of the thesis is as follows:

Chapter 1 discusses the context and background of the issue, the research gap, research aim and objectives. The contribution made by the PhD research to knowledge is also stated in this chapter.

Chapter 2 provides a detailed account of the different energy sources available in Sub-Saharan Africa. Both the renewable and non-renewable energy sources available in the region are discussed in detail using relevant statistics, to provide greater insights into the enormous energy sources in the region. The chapter concludes by pointing out that the account of the energy sources and infrastructure provided in the chapter, suggests that interregional cooperation of trade and supply of energy may be the way forward for the region, especially when considering the uneven distribution of the energy resources, the capital constraints and the low income level across the SSA region.

In Chapter 3, a critical review of the main studies from the theoretical and empirical literature on aggregate and disaggregated energy demand in developing countries is presented. The first section of the chapter starts with coverage of the theoretical underpinnings of the study of energy demand, that is, the neo-classical economic theory of consumers' utility optimising behaviour (Bhattacharyya and Timilsina, 2010; Dramani and Tewari, 2014). This theoretical lens is employed within utility theory and consumer behaviour in a microeconomic context, in the form of a household production function or a utility maximisation function. The section concludes by showing how the microeconomics concept is used as the framework for the analysis at the macro level with mathematical derivations. The chapter then explores the different classifications of energy models in the literature, including static or dynamic, univariate or multivariate, top down or bottom up, identity versus structural or market share based approaches and forecasting models (see, for example, Jebaraj and Iniyan, 2006; Urban *et al.,* 2007; Swan and Ugursal, 2009; Suganthi and Samuel, 2012). The section ends by explaining the top down econometric models and the bottom up engineering and statistical models using relevant examples and studies from the literature. The chapter continues by reviewing different prior studies that analysed the aggregate and/or disaggregated energy consumption-economic growth nexus as well as energy intensity. The chapter concludes by identifying the main variables used for energy demand analysis, the shortcomings of existing studies and the frameworks used in the literature. The identified variables form the basis for the relevant variables to be considered for the empirical model in this PhD study.

Chapter 4 provides an account of the methods used for the estimation of panel data models, which is the main methodological framework used for the empirical analysis in this PhD study. The definitions, derivations and interpretations of panel data models are discussed as adopted for the estimation of the aggregate and disaggregated analysis. The panel methods discussed is grouped under two main headings: panel linear models, and non-stationary linear models.

Under the panel linear models, the fixed effects and random effects models that take into account the unobserved differences in the countries analysed in Sub-Saharan Africa are discussed. The Prais-Winsten (PW) regression model is also explored due to its ability to correct for autocorrelation and heteroscedasticity, if detected. Under the other classification of panel models, the issues of stationarity and unit roots are explained, a basis which is then used for the discussion of panel unit root tests, before moving on to the discussion of panel cointegration. Both methods are used in the data analysis presented in Chapter 6. The chapter concludes with the motivation and rationale for the chosen methodological techniques for the empirical analysis.

Chapter 5 presents the dataset employed for the econometric analysis. In order to construct reliable econometric energy models and obtain valid results from which relevant inferences about energy policies can be drawn, it is important to have consistent long time series data on energy demand and the influencing factors (Pesaran *et al.,* 1998). Therefore, the variables used are described for both the aggregate and the disaggregated energy demand in Sub-Saharan Africa. Sources of each of the variables used in the dataset are given, and the measurement problems associated with them is also discussed.

Chapter 6 presents the results of the estimations obtained using the econometric software STATA 13. The econometric procedure employed for the aggregate energy demand analysis can be divided into three stages. First, the descriptive statistics of the data set used is examined alongside the analysis of the properties of each variable included in the regression through visual inspection of the plots of the relevant series and formal unit root tests. The *a priori* expectations of the relationship between the dependent variable and each of the independent variables are also stated. Second, the cointegration tests and the significance of the long-run estimated coefficients are presented and discussed. Third, the results of the corresponding Error Correction Model (ECM) estimates are illustrated and discussed.

The results of the econometric analysis disaggregated by energy types is based on the determinants of demand for kerosene, petrol (gasoline), liquefied petroleum gas (LPG), solid biomass, diesel and electricity in SSA. The linear panel models in the context of fixed effects, random effects and Prais-Winsten models used in balanced panel models were used for the disaggregated analysis. Attention was paid to the issue of multicollinearity, and the test was carried out in all the models that were analysed and presented.

Chapter 7 offers a more in-depth discussion of the significance of the empirical results obtained. The discussion also reconnects with the findings of previous empirical studies that investigated areas closely linked to the objectives of the present study. It is worth noting, however, that since no other study has attempted to analyse both the driving forces of aggregate energy demand and disaggregated energy demand by fuel types in Sub-Saharan Africa in a single study, a direct comparison of the results with previous empirical work was not straightforward.

Chapter 8 begins with a summary of the main findings of the doctoral research. This is used to draw out various policy implications for SSA, whilst also highlighting the limitations of the research and profitable avenues for future research.

Chapter 2 : Energy Sources in Sub-Saharan Africa

2.1 Chapter overview

This chapter examines the main energy sources in Sub-Saharan Africa (SSA). It starts with a brief overview of the region, followed by the identification and discussion of energy sources. The chapter concludes by identifying the main energy issues affecting SSA.

2.2 Regional overview of Sub-Saharan Africa

Sub-Saharan Africa is located in the Southern part of the Sahara desert, comprising of four main sub-regions, Western, Central, Eastern and Southern Africa (as shown in Figure 2.1). The region is the second largest in terms of landmass after Asia, with 47 countries. It also has the second highest population when compared with other regions of the world. Sub-Saharan Africa occupies a land mass area of 23.6 million (sq.km) and is inhabited by 1.01 billion people (World Development Indicators, 2015).

Approximately three quarter of the countries in the SSA region are among the top 50 poorest countries in the world. In other words, 35 out of the 47 countries in the region are among the 50 poorest countries globally. Common features characterising these countries include: low income level; low production level; poor market structure; unskilled labour; and a high mortality rate. The average annual economic growth rate in the region is about 4%, while the percentage of people that live in poverty in Sub-Saharan Africa is 46.8% (World Development Indicators, 2014).

Furthermore, from the total population in SSA of over a billion people, approximately 60% of them lack access to electricity, with 30 countries having systematic power shortages and/or rationing. The electrification rate and the average annual per capita consumption of power in the region are the lowest when compared to the other regions in the world. For instance, the electrification rate in other developing regions like South Asia and Latin America are 70% and 94% respectively, compared to Sub-Saharan Africa at 32% (IEA, 2013). This is in sharp contrast to the abundant energy sources available in SSA (IEA Africa Energy Outlook, 2014; Onyeji, 2014). Therefore, the need to estimate the energy demand in the region cannot be overstated.

The discussion that follows explores in detail the two main energy categories in SSA: non-renewable and renewable energy sources. However, it should be noted that the energy sources are not equally distributed across the region (Kebede *et al.,* 2010, Mandelli *et al.,* 2014), as will be shown in the discussion that follows.

Figure 2.1: Regional Map of Africa (**Source:** IEA, 2014) This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

2.3 Non-renewable energy sources

Sub-Saharan Africa (SSA) has abundant fossil fuels and uranium resources in different countries across the region. The region is endowed with oil, gas, coal and uranium resources which may be sufficient in meeting energy needs of SSA consumers. A special report by the International Energy Agency on Africa's energy outlook (IEA, 2014) states that the proven oil, gas and coal reserves in the region would be sufficient for another 100, 600 and 400 years, respectively. These estimates include new oil and gas reserves found in the last five years in SSA, which account for 30% of the total new discoveries in the world (IEA, 2014).

Specifically, new discoveries were made in Kenya, Tanzania, Uganda and Mozambique. Examples include the offshore Rovuma basin with a large gas deposit in Mozambique, the Kwanza Basin in Angola with a large deposit of pre-salt oil and gas, and the Keta-Togo-Benin oil Basin in Nigeria. These findings by the IEA clearly suggest two things. The first, is the increased interest in the African oil and gas resources by the international community (market). The second, is the fact that the prevailing energy crisis (for example, low access to electricity) in the region is due to the low use of the available energy resources.

Interestingly, Kebede *et al.* (2010) suggest that three countries in the region, namely South Africa, Namibia and the Republic of Niger, are among the top ten uranium producers in the world. In addition, South Africa is also home to the largest coal reserve in the region. The next subsection provides a more detailed account of the four non-renewable energy sources mentioned in this section.

2.3.1 Fossil fuel

According to the IEA (2014) report, about 7 percent of the total world reserve of crude oil estimated at about 65 billion barrels, is located in the Sub-Saharan Africa region, while the total gas reserve is estimated at 6 percent of the world total. Some studies assert that 13 countries in the region have oil in commercial quantity and are exporters of crude oil (EIA, 2013; KPMG, 2013). These include

Angola, Cameroon, Chad, Congo, Cote D'Ivoire, Democratic Republic of Congo (DRC), Equatorial Guinea, Gabon, Ghana, Niger, Nigeria, South Sudan and Sudan.

Oil production is led by Nigeria and Angola, which might explain why Angola is the second largest exporter of oil to China, after Saudi Arabia (Vines, 2012). However, most of the oil and gas reserves are located in the western regions (Kebede *et al.,* 2010). This is evident by the layout of the Gulf of Guinea which is the main oil basin in the region, with Nigeria having the highest number of reserves. The basin extends from the North-Western part of Angola from Guinea, to Nigeria, Ghana, Ivory Coast, Democratic Republic of Congo, Congo (Brazzaville), Gabon, and Cameroon (KPMG, 2013).

Prior to 2012, Sudan and South Sudan followed Nigeria and Angola in the level of oil production, and thus were the third largest producer of crude oil in the region. Most of the oil reserves are situated in South Sudan which is a land-locked country. On the other hand, Sudan has a port known as the Bashayer port, which stretches across the Red Sea. South Sudan uses the Bashayer port in Sudan for the export of its crude oil to the outside world. The political conflict and oil revenue sharing formula between the two countries, post-independence (2011 onwards), affected the production level and oil exports.

The discovery of the Jubilee field in the Gulf of Guinea in 2007 along the Western part of the shoreline, enabled Ghana to join the league of oil producing countries in Sub-Saharan Africa. The Jubilee field was discovered during the exploration of the deep waters through the drilling of Mahogany-1 by Kosmos Energy (KMPG, 2013). Oil production and exploration in the deep water oil reserve started in 2010 but they are still at a relatively low level when compared to Nigeria or Angola. To exemplify this point, the proven oil reserve in Ghana is 700 million barrels, compared to Nigeria and Angola at 37 and 13 billion barrels, respectively. Nevertheless, Nigeria is the leading oil producer in the region, followed by Angola, with Ghana producing the lowest amount of oil (see Figure 2.2 below). In 2013, the top three oil producers in SSA were Nigeria, Angola and Gabon. Each produced 32 billion barrels, 11 billion barrels and 3.9 billion barrels of oil, respectively.

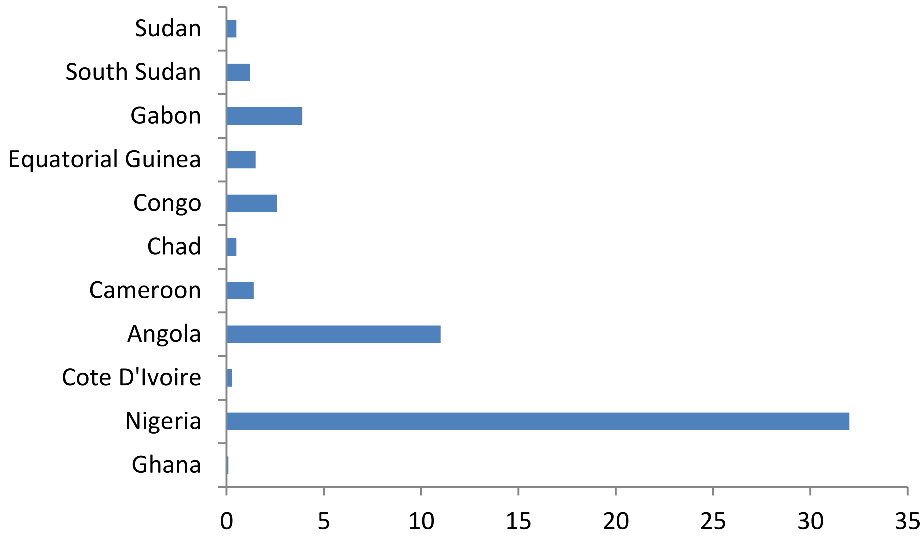

Figure 2.2: Oil Production in 2013

(**Data source:** IEA, 2014)

Shell BP discovered crude oil in 1956 in the Eastern part of Niger Delta in Nigeria, at Oloibiri. Exploration of the resource started in 1958, when 5,100 barrels were produced the day the field came on stream. In the Central African region, oil in Cameroon was discovered in 1955, through the discovery of the Logbaba and Souellaba oil reserves. The commercial potential of the oil was realised in 1972, with the exploration of Betika oil basin. New oil wells were also discovered in Congo and Gabon's deep waters in 2013 (IEA, 2014). Furthermore, the Zifiro complex exploration in the North Western region of Bioko Island in Equatorial Guinea was the beginning of oil production in the late 1900's in the country. Lastly, the construction of the Chad-Cameroon pipeline facilitated oil exports and enabled Chad to start oil production in 2003 (IEA, 2014).

2.3.2 Gas

The Sub-Saharan African region holds about 6% of the world total natural gas resource. The share of the gas reserve in the global percentage is due to an increase of 80% in the new gas reserves discovered within the last few years. The close relationship between gas and oil, has led to several decades of waste of gas through gas flaring in the exploration process of crude oil. Recent awareness about the commercial and environmental impact of gas flaring, notably by regulations from the Global Gas Flaring Reduction Partnership (GGFR), have led to a reduction in the amount of gas flared in recent years. In 2008, 15.5 billion cubic meters (bcm) of the total marketable natural gas of 31.35 bcm in Nigeria was lost to the atmosphere through gas flaring, mainly due to lack of facilities required to capture the gas (Mandelli *et al.,* 2014).

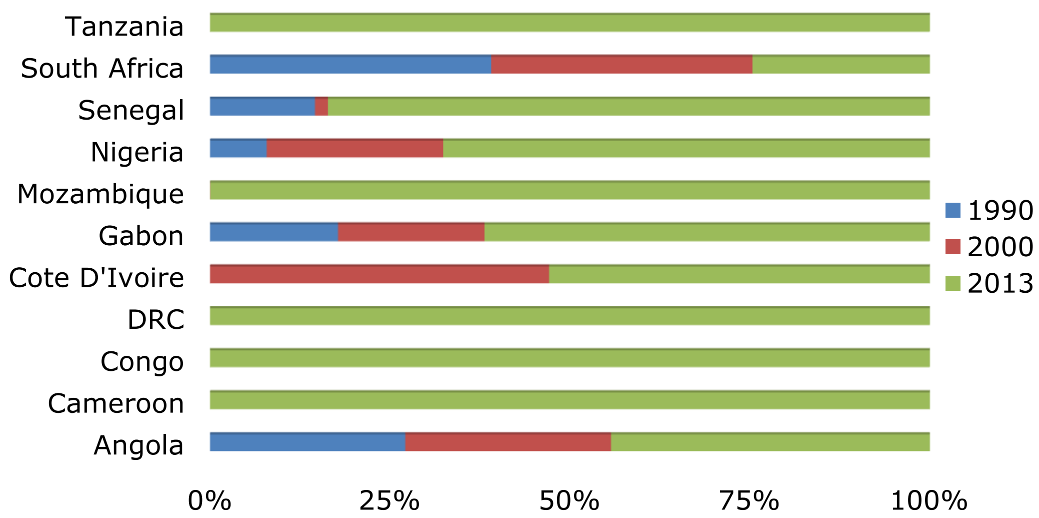

Figure 2.3: Gas Production in SSA in 1990, 2000 and 2013

(**Source:** IEA World Gas Statistics Database, 2015)

From Figure 2.3, it is evident that there has been an increase in gas production in SSA in recent years. In 1990 (blue parts of the chart), only 5 out of 11 countries produced gas. That is, Angola, Gabon, Nigeria, Senegal and South Africa were the only gas producers. After a decade, more gas discoveries and exploration in Cote D'Ivoire and Mozambique with reduction in gas flaring, led to an increase in the total amount of gas produced. In recent years, there has been a noticeable increase in gas production in the region, coupled with a recorded reduction in gas flaring, which can be attributed to increased awareness and regulation. The 2013 (Green section of Figure 2.3) part of the chart shows that, after the year 2000, four more countries joined the gas producing countries, with existing players producing more gas. Cameroon, Congo, Democratic Republic of Congo and

Tanzania joined in 2007, 2003, 2003 and 2004, respectively (IEA World Gas Statistics, 2015).

The drilling of a well in the Pande field by Gulf oil in 1962 led to the discovery of gas in Mozambique. Sasol oil also drilled five wells in the Tamane area and now explores gas in 33 gas wells in the country. This comprises of 15 and 18 production wells in Pande and Tamane respectively (Collings, 2002). Mozambique exports her gas via pipeline to South Africa.

In 2012, Nigeria was among the top five exporter of LNG (Liquefied Natural Gas) in the world, and exported an estimated amount of 950 billion cubic feet of LNG through the country's six LNG trains (EIA, 2013). Nigeria and Equatorial Guinea transports LNG to Asia, Latin America and Europe. Angola started the exportation of LNG in 2013. The other gas producing countries use their production for local gas consumption. In Cameroon, Congo, Cote D'Ivoire, South Africa, and Tanzania, the gas produced was used primarily for power production in industries. Table 2.1 below shows the domestic natural gas production in gas producing countries in the region, between 2009 and 2013. In the table, countries are also ranked on the basis of the total amount of gas produced.

Table 2.1: Five years Natural Gas Statistics

Country	2009	2010	2011	2012	2013	Total	Rank
Angola	690	733	752	760	885	3820	6
Cameroon	315	318	331	346	346	1656	8
Congo	56	103	151	157	157	624	9
DRC	9	9	9	9	9	45	11
Cote D'Ivoire	1518	1655	1632	1780	1780	8365	3
Gabon	248	331	373	384	384	1720	7
Mozambique	3000	3284	3444	3863	4309	17900	2
Nigeria	24409	32540	38341	41201	34425	170916	1
Senegal	20	28	45	46	46	185	10
South Africa	1235	1543	1362	1170	1170	6480	4
Tanzania	655	787	870	995	995	4302	5

(**Data Source:** IEA, World Natural Gas Statistics)

In Table 2.1, data are expressed in million cubic meters. DRC denotes the Democratic Republic of Congo. From the data presented in the table, gas production was led by Nigeria, followed by Mozambique, while the least production was recorded in Democratic Republic of Congo. A visual representation of the data in Figure 2.4 highlights the gas production in the 11 countries, using percentages.

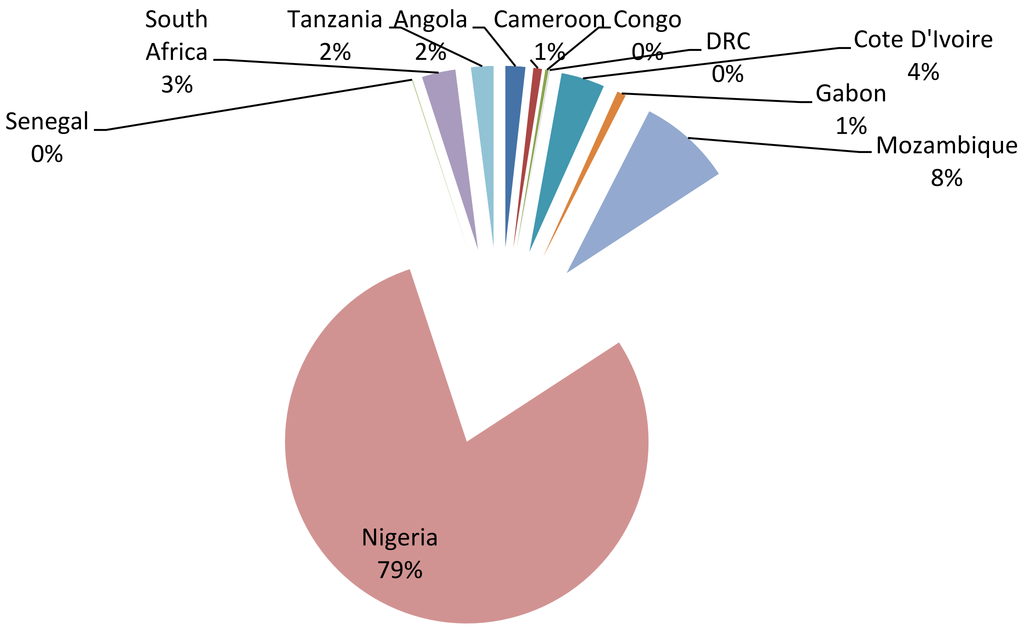

Figure 2.4: Chart Showing Gas Producers in SSA by % share, 2009-2013

2.3.3 Uranium

Three countries in Sub-Saharan Africa, Namibia, South Africa and the Republic of Niger, are among the top ten uranium reserve holders in the world. Studies carried out by the International Energy Agency (IEA, 2014) confirm that the amount of uranium owned accounts for about 18% of the world total uranium sources. Specifically, 16% of the uranium used for nuclear power production is extracted from Namibia and the Republic of Niger (Kebede *et al.,* 2010). This is also supported by Kessides (2014), who points out that in 2011, four countries in

SSA were among the top 15 uranium producers in the world. Apart from the three countries mentioned above, Malawi was included in this ranking.

In Namibia, uranium was discovered in a desert around 1928, while commercial exploration started in 1976. The country has the largest uranium reserves in the region. The three main mines operating in Namibia are Rossing, Trekopje and Langer Heinrich, which are located in the East and Northeast of Swakopmund and West of the Walvis Bay, respectively (WNA, 2014). In 2013, the annual volumes of uranium mined were 2,098, 2,043 and 186 tonnes of U (U for Uranium) at the Langer Heinrich, Rossing and Trekopje mining sites, respectively (WNA, 2014); which according to the International Energy Agency amounted to 8.2% of the total global uranium produced.

The uranium ore in the Republic of Niger is referred to as 'Africa's purest' because of its high quality. The Republic of Niger is Africa's second largest uranium producer. The French Bureau de Recherches Geologiques et Minières (BRGM) discovered uranium during copper exploration in Niger in 1957 while the mining volume of uranium in commercial quantity started in 1971, with the country producing 7.7% of the total mined uranium in the world as at 2013. The three major mines in the country, namely the Somair, Cominak and Somina uranium mines in the mining town of Arit and Akokan, produced a total of 4,528 tonnes of U in 2013 (WNA, 2014).

Uranium was discovered in Malawi in the 1980s, while exploration and production started in 2009. South Africa has the second largest global accessible uranium reserve in the world but the uranium is derived as a by-product with gold or copper (Dasnois, 2012). Other countries in SSA are also considering the inclusion of uranium in their energy mix in order to reduce their dependence on oil. According to the World Nuclear Association, other countries with uranium resources in the region are: Botswana, Central African Republic, Gabon, Guinea, Kenya, Mali, Nigeria, Tanzania, Uganda and Zambia (WNA, 2016).

The Southern African Power Pool (SAPP), which comprises of 12 countries in the Sub-region, with a well-developed interconnection infrastructure, has a total installed capacity of 61,894MW (SAPP 2016 annual report). An estimated 3.5% of the total energy in the SAPP is generated by nuclear energy, with almost 80% of the total capacity in South Africa (SAPP, 2013). Therefore, most countries in the sub-region have nuclear power in their energy mix due to the regional integration of supply. South Africa is the only country among the 47 countries in the SSA region that has operating nuclear reactors located in Koeberg. South African nuclear energy makes up about 5% of the energy mix, representing the only country in the region with two functioning nuclear plants, both located in Koeberg (EIA, 2013).

Nuclear power stations are more cost efficient than fossil fuel plants using oil and gas, due to relative price stability as most of the cost involved lies in building the nuclear plant itself. Construction of nuclear reactors is expensive but it is required in SSA in order to reduce gas or oil shortage problems experienced in most of the electricity plants across the region. There is a need for all sub-region energy blocks in SSA to include nuclear power in their energy mix to bridge the energy demand deficit. Building a nuclear power plant is a very capital intensive project but energy integration and cooperation is thought to reduce the financial burden on individual countries. As part of plans to develop the West Africa Power Pool (WAPP), seven countries in West Africa signed a three-year action plan to jointly develop an integrated West Africa nuclear power station. West African Integrated Nuclear Power Group (WAINPG) was created in 2015 by Benin, Burkina Faso, Mali, Niger, Nigeria and Senegal (WNA, 2016).

Mining of uranium has economic, political, social and environmental impacts which must be considered carefully. First, uranium is the major resource used in building nuclear weapons. This constitutes both social and political problems, and the need for all countries with uranium mines to be part of the Nuclear Non-proliferation Treaty (NPT), to ensure safety and the use of this resource solely in nuclear power plants. Secondly, stringent policies and legislation must be implemented to ensure the safety of miners, local population near the site and careful disposal of waste from the exploration. Lastly, SSA countries must ensure that policies that forbid the unlawful release of Uranium into the atmosphere are enforced. Dasnois (2012) highlighted some of the adverse environmental effects of uranium. The author argues that if released into the atmosphere, it could lead to the deterioration of water quality and a reduction of the water needed for biodiversity and a healthy functioning of the ecosystem along with the loss of habitat and associated threats to some plants and animals in the region. Despite its potential in meeting the region substantial energy demand, careful consideration and planning is required to eliminate the drawbacks associated with using uranium.

2.3.4 Coal

Most of the proven coal reserves are in the Southern part of Sub-Saharan Africa**.** The known 36 billion tonnes reserves are in South Africa, Mozambique, Botswana, Tanzania, Zambia, Swaziland and Malawi. South Africa has 90% of the reserve (IEA, 2014) and, on a global basis, it has the ninth largest recoverable coal deposit (EIA, 2013).

The production of petrol liquids in South Africa is through the highly developed liquefaction of coal at the Secunda plant, in Sasolburg, where the production started in 1955 (Collings, 2002). This involves the production of synthetic gasoline and diesel from coal (CTL) and also, the production of petrol liquids from gas (LTG) at the Mossel Bay plant (EIA, 2014). The plant is one of the largest of such operations in the world (Collings, 2002). The Richards Bay coal terminal is used for coal export, seventy million tons of coal in 2013 were exported, which is about a quarter of the total coal produced that year (EIA, 2014).

A report from the US Energy Information Administration (EIA, 2013) estimates that coal made up 72% of the primary energy consumed in 2013. The report also states that a high reliance on coal makes South Africa the world's 14^{th} largest carbon dioxide emitter and the first among the African countries (EIA, 2014), which is clearly as a result of coal fired stations and plants in the country. Despite their inherent environmental challenges, the use of coal fired stations in South Africa has some advantages over oil and gas fired plants. These include low storage cost, little or low price fluctuations and, above all, lower carbon dioxide emissions than traditional biomass used mainly in SSA. Although Nigeria does not have coal in her energy mix, it is still ranked as the second emitter of CO_2 due to gas flaring.

Mozambique and Zimbabwe are also major exporters of coal in the region. While Mozambique is much under developed, Zimbabwe is still using the traditional low cost open cast method, which has slowed down the full exploration of coal resources (IEA, 2014). Furthermore, coal fired power plants are being installed in Madagascar, Zambia, Botswana and Zimbabwe, using the coal produced locally for the production of electricity.

2.3.5 Hydrocarbon basins and energy infrastructure

The hydrocarbon basins in Sub-Saharan Africa may be classified into five main categories. This comprises of the Niger Delta basin, East African rift, East African coastal region basin, West African transform margin, and West coast pre-salt (Figure 2.5). Studies carried out by the US geological survey (USGS, 2012) found that the Niger Delta basin is the 12th richest but non-fully explored hydrocarbon basin in the world. The Niger Delta basin is composed of several hundreds of small deposits of oil wells, mainly located in the Nigerian water. The Eastern part of the well that is the Niger Delta basin, is the source of Cameroon and Equatorial Guinea oil production. The Eastern Africa coastal regions contain several gas reserves including the Rovuma basin, which is located in the Southern part of Tanzania and in the Northern part of Mozambique.

Exploration in the last five years estimates the basin to have 5 tcm of natural gas (IEA, 2014). West Africa transform margin extends from Mauritania to the Niger Delta region. This includes the Jubilee oil field of Ghana, and Liberia, Sierra

Leone and Cote D'Ivoire's oil reserves. However, the commercialisation of the oil resources in Liberia, Sierra Leone and Cote D'Ivoire are yet to be explored.

The discovery of the East African rift hydrocarbon basin changed the oil and gas industry traditional players in Sub-Saharan Africa, by including the East African countries in the production frontier (Vines, 2012). Specifically, in 2007, the KingFisher oil basin, with an estimated 1.7 billion barrels of oil was discovered in Uganda, also the Lokichar basin with about 600 million barrels of oil was discovered in Kenya, while the Ogaden basin was discovered in Ethiopia.

The main oil and gas infrastructures in the sub-regions include the West African Gas Pipeline (WAGP) in four Western Africa countries, Pande and Tamane gas production wells pipeline from Mozambique to South Africa Secunda plant, and the Lamu Port South Sudan Ethiopia Transport Corridor (LAPSSET) deep-water pipeline under construction.

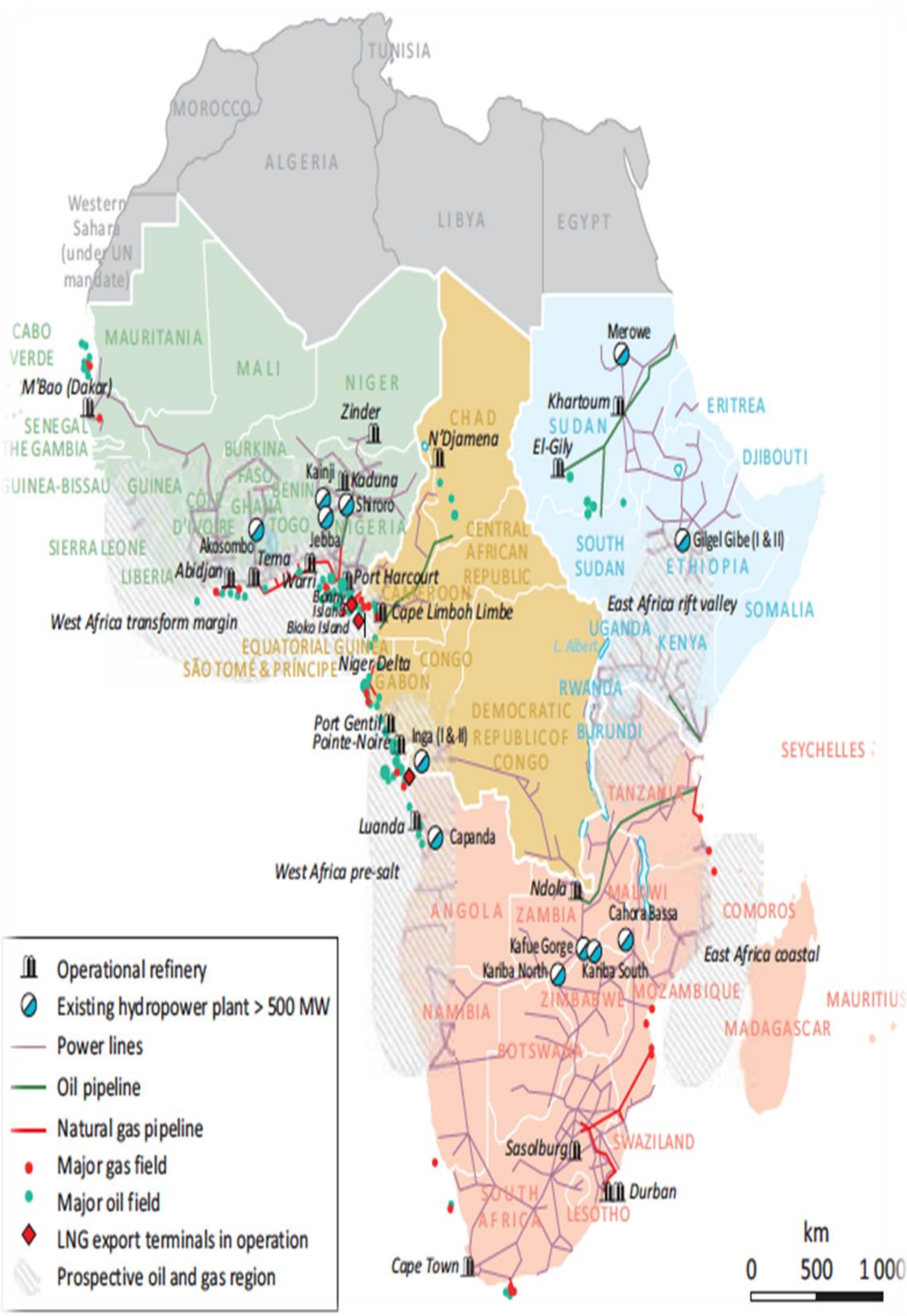

Figure 2.5: Hydrocarbon Basin and Infrastructure (**Source:** IEA, 2014) This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The West Africa Gas Pipeline is one of the most notable energy infrastructures in the region; the 678 km pipeline which runs from Nigeria to Benin, Togo and Ghana was completed in 2007. The pipeline, which was opened in 2010, is known as the first high pressure gas transmission network in Africa (Kessides, 2014). The governments involved in the pipeline construction aim to achieve a cleaner source of power generation and to provide long-term access to modern energy in the region (WAGPCO, 2013). The gas is to be supplied from Nigeria to three countries' thermal power station plants in addition to Nigeria's own. Regional integration in the supply of purified natural gas in the four West African countries through the West African Gas Pipeline (WAGP) suggests that this could be a way forward in solving the energy crisis in the region. With only four countries connected to the pipeline out of the eighteen countries in the sub region, this creates opportunities for more intra-regional trade in the sub-region and at a continental level (studies that support this view include Guansounou *et al.,* 2007; Iwayemi, 2008; Ouedraogo, 2013; Demierre *et al*., 2015; Uddin and Taplin, 2015).

The pipeline used for the export of Mozambique's gas to South Africa is an integrated 865km pipeline network that runs from the Pande and Tamane gas production wells in Mozambique to South Africa's Secunda plant (Collings, 2013). The pipeline according to one of the reports of the Sasol petrochemical company, which operates the coal liquefaction plant in South Africa, will enable the production of a cleaner energy through gas, with less environmental pollution associated with the production of synthetic fuels from coal in South Africa (SASOL, 2012). It will also provide more resources to be used in the production of petrochemicals and green power in South Africa.

The ongoing Lamu Port South Sudan Ethiopia Transport Corridor (LAPSSET), whose construction began in 2012, is a deep-water port in Northern Kenya (Vines, 2012). The pipeline will integrate oil production and transport oil between Kenya, South Sudan and Ethiopia, through regional pipelines, railways and roads in the countries. Vines (2012) also suggest that the revenue from the oil and gas in the region should be used in building more regional infrastructure.

From the discussion above, most of the energy infrastructure seems to be suitable for the trade or transport of non-renewable energy sources (see, for example, oil and gas pipelines, and power lines in Figure 2.5), in particular for oil and gas distribution in the region. The presence of more fuel based infrastructure may be explained, for example, by the fact that most of the power generation in the region is produced from fossil fuel. In 2011, over 80% of the regional power generated emerged from fossil fuels (Kessides, 2014). The main power pools are Southern Africa Power Pool (SAPP), East African Power Pool (EAPP), West African Power Pool (WAPP) and the Central African Power Pool (CAPP).

Despite the existence of the above pools, there is still a need to provide for additional inter-regional power pool to cater mainly for renewable energy sources, to make some projects (such as the Grand Inga Dam proposed to be built in the Democratic Republic of Congo) economically viable. This will require extended cross border distribution and electricity transmission links between the home and host domestic economy, because it is expected that the power generated will be more than what is needed in the host domestic economy. According to Kessides (2014), the resultant effect of more inter regional connections are technical and political constraints. The construction of long electricity transmission cables in the region requires high technical know-how and skills, in terms of engineering and construction skills, and may also be affected by some political factors like conflicts and poor land tenure systems (Kessides, 2014).

2.4 Renewable energy sources

The sections above focused on the non-renewable energy sources, hydrocarbon basins and the energy infrastructure in Sub-Saharan Africa. The sub sections that follow highlight the renewable energy sources in the region in more detail. Sub-Saharan Africa is not lacking in renewable energy sources, which could be exploited to meet the energy demand in the region. As argued by Onyeji (2014), the available energy resources are in abundance and could be used to meet both the present and the future energy needs in the region. Each of the renewable energy sources available in SSA will be discussed in detail below, notably: hydropower, solar power, wind power, geothermal sources and bioenergy.

2.4.1 Hydropower

The river systems in Africa are among the largest in the world and account for about 13% of the total hydropower potential available globally. As noted by Kebede *et al.* (2010), five of Africa's rivers - the Nile, the Congo, the Niger, the Volta, and the Zambezi - form the largest river systems in the world. Specifically, all the rivers mentioned are located in Sub-Saharan Africa, except the Nile which passes through Egypt to empty into the Mediterranean Sea. Furthermore, most of the hydropower potentials are in Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Nigeria, Guinea, Senegal, Angola and Mozambique (IEA, 2014).

Hydropower ranks top among the different renewable energy sources used in Sub-Saharan Africa, and its percentage in the total energy mix in the region is estimated to be about 16% (Onyeji and Jones, 2014). The river system in the Democratic Republic of Congo, known as the Congo, is the second largest river in the world after the Amazon, with a 42,000m^3/s (meter cubic per second) flow speed, and is also the second biggest in Africa after the River Nile, covering a land area of 4,700 kilometers (Onyeji and Jones, 2014). However, most of the hydropower potential available in the river has not been tapped to meet the energy demand in the region, due to financial constraints and conflicts in the country. This issue is at the forefront of discussions among policymakers in the region, including the New Partnership for Africa's Development (NEPAD).

A consensus was reached among major stakeholders to use the available hydropower potential to meet the power needs of the consumers in Sub-Saharan Africa. Stakeholders include governments in the region, national agencies (SAPP, NEPAD) and international organisations like the World Energy Council, World Bank, Africa Development Bank and the European Investment Bank. Propositions by the major stakeholders led to the proposed lnga lll (4.8GW) and Grand lnga project of 44GW, in the Western side of the Democratic Republic of Congo, which will be implemented in seven phases. However, it should be borne in mind that as with any river system, the river is subject to seasonal variation on a year-to-year basis. Also, building the large hydro plant will require more developed interconnection of energy supply systems and power pool in the region, as the power generated will be too large (estimated at 40,000 MW) for domestic supply and will need to be exported to other parts of the continent (Onjeyi, 2014). Hence, the infrastructure and projects that will facilitate the expansion of interconnections among the sub-regions should be taken seriously (Kalitsi, 2006).

Hydropower is estimated to have the cheapest on-grid cost of generating electricity at $55/ MWh among the different renewable sources, based on different studies (see on-grid large hydro cost in Figure 2.6). Moreover, as argued by Kalitsi (2006), building dams and hydro projects have the advantage of facilitating water supply, irrigation systems, fishing activities and could also be a good source of tourist attraction where they are built. As stated earlier, there are many unused sources of hydro potential in Sub-Saharan Africa. The total estimated hydropower potential in the region is 283 GW, out of which only about 10% has been used so far. For instance, in 2012, the total on grid power generation capacity in the region was 90GW, which comprises of 45% coal mainly from South Africa, 22% hydro, 17% gas from Nigeria, while 14% was from oil (IEA, 2014). Clearly, if more hydropower is exploited and employed to provide more power, this will reduce the prevailing energy poverty in the region.

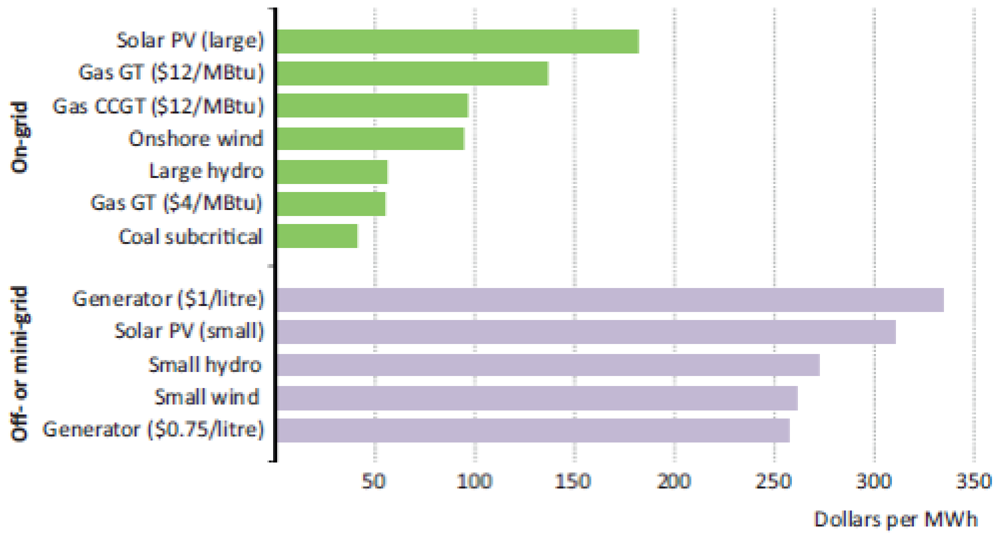

Figure 2.6: Comparative cost of on-grid and off-grid power generation sources
(**Source:** IEA Africa Energy Outlook, 2014)

Apart from the low cost of production associated with using hydropower for energy generation, there are additional advantages that are associated with using it as a means of energy supply. This was highlighted in the study by Oyedepo (2012), who suggested additional benefits such as reduced impact on the environment, the potential of constructing a fishery, irrigation systems, flood prevention and, of course, power generation.

Figure 2.6 above, presents the cost of different on-grid and off-grid power generation sources in SSA. Coal has the cheapest cost, while the cost of plant using gas varies depending on the specification. Wind has the second lowest cost, while solar is the most expensive out of the renewable on-grid sources of generating electricity. The lower part of the chart highlights the cost of off-grid power solutions.

2.4.2 Solar power

Sub-Saharan Africa has one of the best solar energy potential in the world. This stems from a geographical location near the equator, which provides the region 320 days of sunshine (in most countries) per annum. The irradiance level varies across sub-regions, with the Southern sub-region having the highest insolation level of 2,500KWh per m^2 per day (IEA, 2014). However, in general, the whole region has an insolation level of about 2,000 KWh per m^2 per day in a year (IEA, 2014). This suggests that most part of the region meets the recommended threshold of 5 KWh/m^2/day needed for the efficient functioning of solar thermal facilities (Deichmann *et al.,* 2010). This is also supported by Onyeji and Jones (2014), who argued that the average solar power intensity between the range of 3,000 to 7,000 $W/h/m^2$ in Sub-Saharan Africa, is higher than what is needed normally to supply domestic loads.

The power cost for a large on-grid and small off-grid solar PV technology, estimated to be approximately US \$180/MWh and \$310/MWh (IEA, 2014), represents the highest cost of producing electricity among the renewable energy sources. The relatively higher cost of production, however, has not discouraged some of the governments in the region. For instance, there is a 155MW Nzema plant in Ghana. Furthermore, photovoltaic panels provide a solution to the low electrification level in most of the countries in the region, by enabling the use of lighting, simple gadgets operation, rooftop systems in homes, in both the rural and urban areas. This particularly suits the off grid market in most of the countries in the region, as there is abundant sunlight to power the systems. Most countries across the SSA region encourage the use of small off-grid solar panels, but there is still a need for governments to make it more affordable through appropriate incentives and subsidies. Especially for the rural off-grid market, as this would also boost socio-economic development and reduce rural-urban migration.

2.4.3 Wind energy

Wind resources vary widely across Sub-Saharan Africa. Depending mainly on the country's geographical location, each country has different on-shore and offshore potential. To be specific, most of the wind power potential lies in the Horn of Africa, the Eastern part of Kenya, the side of the West and Central Africa located around the borders of the Sahara, and lastly some areas in the Southern part of the region (Mukasa *et al.*, 2015). A feasibility study on the Sub-Saharan Africa wind power potential by the African Development Bank (AfDB, 2012), found that

Somalia has the highest potential, followed by Sudan, Mauritania, Madagascar, Chad and Kenya, all with a high on-shore wind power. The study also highlighted that Namibia, Mozambique, Tanzania, Angola and South Africa have large offshore wind energy potential in the region.

The result of the feasibility study, is supported by an earlier study by the African Development Bank Group (2004), as cited by Mukasa *et al.* (2013) who using a Wind Energy Simulation Toolkit (WEST) found that wind energy is highest in countries in the coastal areas, with wind speeds of about 7.0 m/s and 7.5 m/s. Specifically, the coastal area countries identified in the study are Eritrea, Seychelles, Somalia, Cape Verde, South Africa, and Lesotho. The wind energy in these countries could be used for commercial electric power generation (Mukasa *et al.,* 2013), by building a wind power station, which is built in places with a high wind speed by putting in place wind farms.

The exploitable wind power in SSA is about 1,300 GW (Mandelli *et al.,* 2014). However, very little of the available wind power has been used. This is evident by the installed capacity at the end of 2013, which was about 190 MW. Most of the existing wind turbines are located in Namibia, Mauritius, South Africa, Kenya and Eritrea (Mukasa *et al.,* 2013). In 2013, the smaller windmills in the Assab wind park in Eritrea, Mozambique and Namibia generated about 1MW, 300kW and 200kW of energy in their energy mix, respectively.

2.4.4 Geothermal sources

SSA has one of the best geothermal potential in the world, which is mostly concentrated in the Eastern part along the rift valley, with an estimated potential between 10GW and 15GW (IEA, 2014). Two countries in the Eastern region, Kenya and Ethiopia, have the best geothermal sources in SSA, and Kenya being the first country in the region to use geothermal for power generation. The exploration of geothermal sources started in the 1960's through the drilling of two wells in the Olkaria area of Kenya.

Further, the establishment of the Geothermal Development Company (GDC) in Nairobi, the capital city of Kenya, by the Kenyan government has the aim of harnessing the country's huge geothermal potential, and producing energy from the cleaner source. Studies by the International Energy Agency, estimate that geothermal added about 250MW to the energy mix in Kenya, and with the annual drilling of forty wells per annum in the country, another 280MW will be added in the next two to three years (IEA, 2014).

The government in Ethiopia, the Power Africa Initiative and other investors are all working to harness the available geothermal sources to produce power in the country. To illustrate this, the proposed Corbetti power project which aims to generate about 1GW is currently under development and planning, to be built in different phases over the next ten years (IEA, 2014).

Geothermal as an energy source has some advantages over other renewable energy sources. Firstly, it is not seasonal nor fluctuates at different times like other renewable sources with annual variations in supply, like hydro power or solar power sources which cannot be available 24 hours a day, or all year round. It provides a reliable source of energy due to its non-intermittent nature. Secondly, with its low emission rate, geothermal is considered to be a clean source for energy production. Lastly, geothermal could also be used to reduce the energy poverty in SSA, as it provides a good alternative to using oil fired plants which could be expensive and unpredictable in terms of cost due to high dependence on oil which is susceptible to fluctuations in price.

2.4.5 Bioenergy

Several studies support the view that solid biomass is the dominant domestic fuel for cooking, drying and heating in SSA, and is also a key renewable energy source in the region (Kebede *et al.,* 2010; Brew-Hammond, 2007; Karekezi, 2002). This may be explained by the high use of traditional biomass by the majority of the population in the residential sector, and also the availability of forest resources. Indeed, 33% of the landmass in SSA is dominated by forest, illustrating the dependence on solid biomass by countries in the region. Wicke *et al.* (2011) argued that the primary energy supply in most of SSA is composed of 70- 90% of biomass, while at the same time the total energy consumption in these countries is made up of up to 95% of biomass. Furthermore, IEA (2014) estimates that there are about 130 billion tonnes of forest biomass in the region. Most of the forest regions are in Central and Southern Africa, with the largest share located in the Congo basin area. This may explain why traditional biomass from the use of firewood, animal dung and waste, make up a large percentage of the energy mix in most of the countries, as mentioned earlier.

Previous studies such as Owen *et al.* (2013) have noted the benefits of biomass energy for SSA. The benefits of the consumption of biomass highlighted by the study are: large availability; familiarity and lower price than most fossil fuels and electricity sources; job creation, energy security and diversity; reduced climate change; avenue for technological advancement and modernity; commercial investment in electricity generation from combined heat and power (CHP) plant using biomass as the main fuel; amongst others. Jobs can be created with the construction of biofuel plants, which will also help to provide a clean source of energy. The authors argued that biomass in form of charcoal and firewood can be bought according to household income, and its supply is never scarce unlike liquefied petroleum gas. They explained further that when wood fuels are sourced in a sustainable way, they are carbon-neutral and can help to mitigate climate change. This suggests that the adoption of technology for biomass creates advancement for both domestic and commercial application (Owen *et al.*, 2013).

As further pointed out by Onyeji (2014), the use of biofuels and biogas has some advantages. First, the fuel from biomass residues could be used for operating cooking appliances. Second, it reduces the indoor air pollution and green house gas (GHG) emissions associated with the traditional use of biomass through burning, resulting in improvements in health conditions. The time used in gathering the biomass residues especially firewood, animal waste and other agricultural residues could be spent on other productive activities. Lastly, in the rural areas, both bioethanol and biodiesel could be used as cost efficient alternatives for use in vehicles used for transport purposes and the running of some of the agricultural equipments in the farms.

Wicke *et al.* (2011) add that the decrease in the use of traditional biomass could be linked to the increase in income levels in some countries, which shows clearly that the problems associated with the use of this form of energy, could be as a result of the widespread poverty in SSA. For example, according to a report by Mundi (2014), 60.5% of the population in Zambia lives below the poverty line (that is, below $2 per day), which may also explain why most of the people in the rural areas use biomass for cooking, while at the same time about 90% of the household in Zambia use charcoal for cooking. This suggests that both variables are closely linked. For example, in Nigeria, in 2011, traditional biomass made up 83% of the primary energy consumed (EIA, 2013), while the poverty figure stood at more than 70% of the total population, in spite of the increase in economic growth (see appendix 1).

Further, in countries like Burundi, Rwanda, Mozambique, Burkina Faso, Benin, Niger and Madagascar, the amount of traditional biomass in their energy mix ranges between 85 and 91% (UNECA, 2006 as cited in Dasappa, 2011). Recent data also confirms the high use of biomass in SSA as depicted in Figure 2.7 below. In 2013, traditional biomass accounted for most of the total energy consumed: 96% in Angola, Benin, Gabon, Mozambique and Togo; 97% in Cameroon, Congo, Cote D'Ivoire, Eritrea, Kenya and Zimbabwe; 99% in Nigeria and Democratic Republic of Congo; 98% in Niger; 83% in Bostwana and Cameroon; 89% in Senegal; 94% in Sudan and Zambia. Namibia had the least proportion of solid biomass in her energy mix at 52%. Figure 2.7 below, illustrates the data more clearly.

It is evident from Figure 2.7 that energy supply and economic growth have both been outpaced by a higher increase in population size. This may explain why the use of traditional biomass has increased in some SSA countries over time. For instance, the amount of solid biomass in Niger energy mix was 85% in 2006 but the proportion increased to 98% in 2013 due to an increase in the population growth rate. Niger population increased by 30.35% between 2006 and 2013 (World Development Indicators, 2015). The same trend is observed in other countries with high biomass consumption.

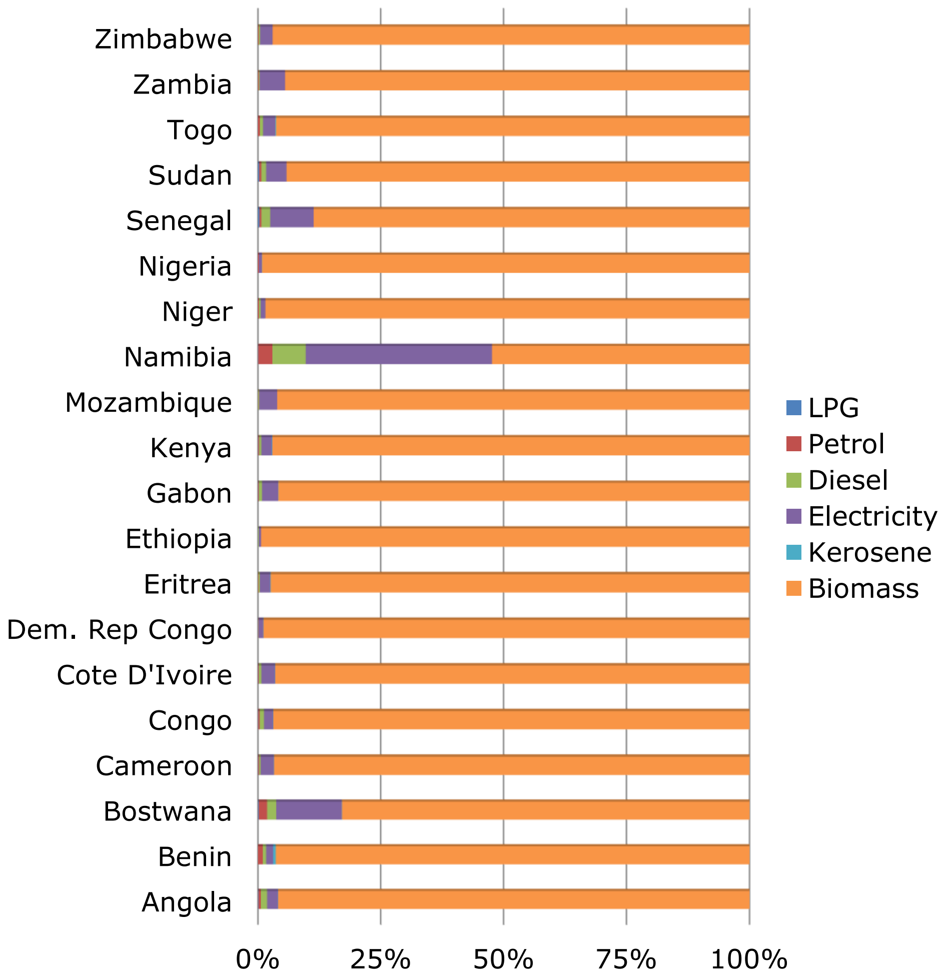

Figure 2.7: Energy Consumption by Energy Types in 2013

(**Source:** IEA Energy Consumption Data, 2015)

However, despite the wide use of traditional biomass and the high level of poverty (above 50%, World Bank, 2014) in most of the countries in the region, some of the countries in SSA are adopting the use of bioenergy systems as part of their energy supply mix. To exemplify this, in the Eastern and Southern Africa region, the installed existing energy from bioenergy is around 325MW. This needs careful consideration, as some studies have shown that energy from bioenergy sources are more expensive and less competitive when compared to those produced from fossil fuel (see, e.g., IEA, 2014). However, this is to be expected as most renewable energy sources are more expensive than using fossil fuel like coal for power generation.

Clearly, at the regional level, there are opportunities to reduce the energy poverty and the high reliance on the traditional use of biomass but challenges and limitations exist.

2.5 Chapter summary

The objective of this chapter was to highlight the main energy sources in SSA, and to discuss each of the available sources in more detail so as to provide a context for this PhD study, and present the energy outlook for the region under investigation. From the previous studies reviewed, and also using reports and data from energy institutions, it is evident that Sub-Saharan Africa has abundant energy resources that can be used to meet the energy needs of consumers. The discussion of the non-renewable and renewable energy sources reveals that both energy sources are available in considerable quantities but they are not evenly distributed across the region. Moreover, the account of the energy sources and infrastructure presented in this chapter clearly shows that interregional cooperation in trade and the supply of energy presents a way forward for the region, considering the capital constraints and the low income level in the region. The next chapter presents a critical review of the theoretical and empirical literature on energy demand.

Chapter 3 : A Critical Review of the Literature on Energy Demand

3.1 Chapter overview

A considerable amount of studies have been published on both the aggregate and disaggregated energy demand in the energy literature. Several attempts have been made at estimating energy demand, with researchers using various methods and obtaining different results, which have either confirmed or contradicted earlier findings. The empirical model used in this study is based on the knowledge gained from the review of the existing literature on energy demand. Specifically, various models which analyse energy demand in developing regions are reviewed and used to develop a comprehensive framework for the empirical analysis in this study.

The objective of this chapter is to undertake a systematic review of the established theoretical and empirical studies on energy demand according to the techniques used for the analysis, in order to provide a framework for the analysis of energy demand in Sub-Saharan Africa (SSA) that accounts and addresses gaps and shortcomings of previous research in this field.

Energy studies, in both developed and developing countries, may be classified into four main groups:

- Aggregate demand for energy
- Disaggregated demand for energy
- Energy consumption and economic growth
- Determinants of energy intensity

Previous findings based on the classification above will be reviewed critically in this chapter. However, before the studies are discussed, the next section provides a discussion of the general theoretical framework underpinning the study of energy demand.

3.2 Theoretical background: Energy demand

The main economic theory used by researchers for the study of energy demand, is based on the neo-classical economic theory of consumers utility optimising behaviour (Bhattacharyya and Timilsina, 2010; Dramani and Tewari, 2014). This is employed within utility theory and consumer behavior in a microeconomic context, in the form of a household production function, or a utility maximisation function. Researchers that studied either the industrial or commercial energy demand have analysed the demand in each sector within the theory of the firm as well (Bhattacharyya and Timilsina, 2010). The microeconomics concept that is used as the framework for the analysis at the macro level is discussed below.

Central to the whole discipline of economics, is the investigation of how prices and income impact on the demand for goods and services. To this end, economic theory has provided some (now well established) theoretical models for such analysis (Lawler and Seddighi, 1987). This is the microeconomics basis for the total energy demand analysis. As pointed out by Lawler and Seddighi (1987, p.1**),** '*the establishment of the law of demand is the main objective of the theoretical frameworks in economic modelling'*. Furthermore, Samuelson and Nordhaus (1992) argued that the law of demand states that when the price of a good increases, the quantity of the commodity consumers will be willing to buy will reduce. Similarly, buyers will be willing to buy more of a good when the price is reduced. This suggests that there is a direct link between the market price and the quantity demanded of a commodity, *ceteris paribus*. The negative relationship between the price of a good and the quantity demanded is known as the demand curve (Samuelson and Nordhaus, 1992).

This concept finds its microeconomic underpinnings in household utility maximisation theory. The demand curve is explained using the concept of utility and consumer behaviour, which can then be used to explain the total demand in an economy (Samuelson and Nordhaus, 1992), at both the micro and macro level (Mensah, 2014). On one hand, the theory of consumer behaviour states that the household tries to maximise utility by dividing the limited money income between different goods and services. That is, the consumer spends his or her income in such a way that the highest level of satisfaction is derived from the consumption of both energy and non-energy goods.

The definition of consumer behaviour given above also suggests that the theory would be useful in analysing how consumers respond to changes in income and price, so as to enable them to derive maximum satisfaction (utility) given the budget constraint at different times (Gould and Lazear, 1989).

According to a definition provided by Samuelson and Nordhaus (1992, p.172)**,**
'*utility is a scientific construct that economists use to understand how rational consumers divide their limited resources among the commodities that provide them with satisfaction'*.

In essence, utility is a common denominator used by individuals or households to compare the satisfaction obtainable between different commodities (Lawler and Seddighi, 1987). This logic assumes that consumers are rational, and know the available preference sets. The preference ordering may be represented using a utility function, and such consumer will choose the most preferred set (Bhattacharyya and Timilsina, 2010). Following consumer theory, utility is measured in an ordinal and not a cardinal sense, suggesting that the numerical value assigned to the utility function is not important provided the ranking of the goods and the consumer preferences are represented (Samuelson and Nordhaus, 1992).

The assumptions underlying the key concepts unpacked in the paragraphs above were deduced by Alfred Marshal in his 'Principles of Economics' (1890) in which they were used to derive the law of demand, later named as 'Mashallian Demand Theory'. Marshall (1890) sets out nine assumptions as follows: commodities have a fixed price; the budget is limited to the money income received by the consumer; there is perfect information about the commodities available; the use of goods and services gives the consumer satisfaction; the satisfaction obtainable is dependent on the quantity of goods and services consumed; the satisfaction derived from the use of goods and services is measured in a cardinal sense; consumers are rational decision makers; diminishing marginal utility is associated with the consumption of commodities and the total satisfaction derived by the consumers from a set of goods or services can be added together to know the total utility derived by the consumer.

As mentioned earlier, the concept of utility is represented using a utility function, to provide an objective function for the consumer, which in turn is used for utility maximisation in decision making problems (Lawler and Seddighi, 1987). Therefore, a utility function represents the level of utility or satisfaction derived from the bundle of commodities; a great tool to be used for the analysis of consumer behaviour (Gould and Lazear, 1989).

Based on the micro analysis, the household energy demand is formulated and solved mathematically using the utility function written below (3.01), to represents a consumer or a household consumption utility function, based on the micro analysis. The consumption of energy is not a 'direct demand' deriving from the demand of services provided by energy, suggesting that it is a 'derived demand'. In other words, energy is always used together with other machineries, household appliances as well as other consumer durables (Pesaran *et al.*, 1998).

First, the utility function of a consumer is represented by:

$$U = U\ (X_1, X_2, X_3, \ldots \ldots \ldots \ldots \ldots \ldots, X_n) \qquad (3.01)$$

The budget constraint of the consumer can be written as

$$Y = P_1X_1 + P_2X_2 + P_3X_3 + \cdots \ldots \ldots \ldots \ldots \ldots \ldots + P_nX_n \qquad (3.02)$$

Next, the Lagrange (L) is set so as to maximise the consumer's utility subject to the budget constraint:

$$L = U(X_1, X_2, X_3, \ldots \ldots, X_n) - \lambda\ (Y - (P_1X_1 + P_2X_2 + P_3X_3 + \cdots \ldots \ldots + P_nX_n)) \qquad (3.03)$$

The partial derivatives of L are set with respect to $X_1, X_2, X_3, \ldots \ldots \ldots \ldots \ldots \ldots, X_n$, while λ is set to zero in all the equations to obtain the necessary conditions

$$\partial l/\partial X_1 = \partial U/\partial X_1 - \lambda P_1 = 0$$
$$\partial l/\partial X_2 = \partial U/\partial X_2 - \lambda P_2 = 0$$
$$\partial l/\partial X_3 = \partial U/\partial X_3 - \lambda P_3 = 0$$
$$\vdots$$
$$\vdots$$
$$\partial l/\partial X_n = \partial U/\partial X_n - \lambda P_n = 0$$
$$\partial l/\partial \lambda = Y - P_1X_1 + P_2X_2 + P_3X_3 + \cdots \ldots \ldots .. + P_nX_n = 0 \qquad (3.04)$$

To solve the equations above, one can equate them to one another and solve for lambda, to get the marginal rate of substitution (*MRS*), i.e., the marginal rate of substitution of one commodity for the other

$$\partial l/\partial X_1/\partial l/\partial X_2 = P_1/P_2 \qquad \cong MRS = price\ ratio$$

$$\lambda = \frac{\partial U/\partial X_1}{P_1} = \frac{\partial U/\partial X_2}{P_2} = \cdots \ldots \ldots \ldots . = \frac{\partial U/\partial X_n}{P_n} \qquad (3.05)$$

The last step is to derive the demand function for energy which is a function of price and income. The assumption here is that both price and income are homogenous of degree zero, and the necessary condition is solved for each equation. The resulting demand equation is represented by the equations:

$$X_1 = d_1(P_1, P_2, P_3, \dots \dots \dots \dots, P_n, Y)$$
$$X_2 = d_2(P_1, P_2, P_3, \dots \dots \dots \dots, P_n, Y)$$
$$\vdots$$
$$X_n = d_n(P_1, P_2, P_3, \dots \dots \dots \dots, P_n, Y) \qquad (3.06)$$

Equation (3.06) signifies that the demand for energy is a function of both the relative prices and real income of the consumer.

However, for use in empirical modelling, other factors that are identified to influence the demand for energy are included, which makes the empirical model to be in the form presented in equation (3.07) below. Where the vector N, represents other driving factors or forces of energy demand.

$$X_n = d_n(P_1, P_2, P_3, \dots \dots \dots \dots, P_n, Y, N) \qquad (3.07)$$

In addition, it is important to discuss how the production sector, which includes industries and firms within them, demands energy as an input in the production process, and achieves the objectives of total cost minimisation and profit maximisation. This is traditionally based on the theory of producers, in the microeconomics branch of economics. The main idea of how the objective is achieved, is based on three main assumptions, as pointed out by Bhattacharyya and Timilsina (2010): (i) a technical limitation to the amount of goods that could be produced in the production process; (ii) machinery cannot produce beyond certain limits at any given period of time; (iii) the inputs used in the production process may not constitute a surplus, or they may be limited (Bhattacharyya and Timilsina, 2010). Having stated the three assumptions, similarly to the household utility maximisation, the firm maximises profits at any given level of production, at the point where the rate of technical substitution is equal to the ratio of input prices.

The account above is intended to provide the theoretical background of the study based on the micro level analysis, which as it is argued by Mensah (2014) can be further extended to analyse energy demand for the entire population (or the whole economy) at the macro level. This could be achieved by treating the whole population under study, like the representative household used in the example above, that takes the consumption decision of energy and non-energy goods, while considering income, prices and other exogenous factors.

To conclude, it is reasonable to suggest that the task of the researcher is to provide a suitable theoretical framework which explains the data from actual markets based on empirical evidence drawing from the theoretical underpinnings covered above. It can be further stated that studies that have analysed the energy demand in a specific sector of the economy, like the residential sector, have also used household production theory, on the assumption that the household produces energy services such as cooling, warming or lighting through the combination of electricity with electrical appliances (see, for example, Blazquez *et al.*, 2013).

3.3 Energy demand models

This section provides a review of some of the models used for the analysis of energy demand. Energy models could also be used for the forecast of energy demand through the use of macroeconomic variables, whilst the results from the analysis are used for managing energy from the demand side, for example, by using them in the planning and drafting of policies aimed at those objectives (Suganthi and Samuel, 2012). However, the modelling of energy demand is not as straightforward when compared to the empirical models used for the analysis of other goods and commodities. This is due to the derived nature of the demand for energy (Pesaran *et al.,* 1998), as explained earlier in section 3.2.

Energy models in the energy literature have been classified in several ways, such as static or dynamic, univariate or multivariate, top down or bottom up, identity versus structural or market share based approaches, forecasting models (see, for example, Jebaraj and Iniyan, 2006; Urban *et al.,* 2007; Swan and Ugursal, 2009;

Suganthi and Samuel, 2012). Furthermore, Pesaran *et al.,* (1998) classified energy models into three main classes. The authors argued that the structural engineering approach, end-use approach or the econometric approach provide the best frameworks for the estimation of energy demand and associated analysis.

Swan and Ugursal (2009) provide a comprehensive review of the different models used in the analysis of residential sector energy consumption by different studies in the energy literature. They point out that two main model-categories are used, the top-down and the bottom up approach. The authors give an account of each technique by highlighting the strengths, objectives and the drawbacks of each model. Furthermore, they assert that the models require different information levels, implementation techniques, and also produce different results which could be used in different situations.

Other researchers, including Torriti (2014), have also attempted to classify energy models used for the residential sector electricity demand, according to the discipline of the researcher and the datasets used for the analysis of energy demand. Such classifications include energy econometricians using aggregate macroeconomic data and electrical engineers using actual or simulated end-user data.

Until now, research on energy demand has focused more on developed countries rather than developing countries, mainly due to low skills and data constraints (Endresen, 2004). Therefore, it is crucial for the researcher to assess the suitability of the model used, to ensure it captures the specific energy system characteristics of the developing countries context. As argued by Urban *et al.* (2007), the economies and the energy systems in developing countries vary from those in developed countries and should be modelled in a different way. This has become even more important in the last few decades due to the increased participation of some of the developing countries resulting from globalisation and greater economic integration, especially China in the global energy scene (Bhattacharyya and Timilsina, 2010).

Researchers including Shukla (1995), Pandey (2002), Urban *et al.* (2007) and Bhattacharyya and Timilsina (2010) hold the view that energy models intended for developing countries should be built taking into account the developing region's peculiar characteristics rather than just adapting models that have been used for developed countries. The authors point out that the common factors in the countries in the developing world include: the prevalence of the use of traditional energy sources like biomass; a large number of informal energy sectors; the existence of a high level of inequality and poverty; structural changes from traditional to a modern way of living; high demand/supply deficit; limited capital; and a low rate of technological diffusion.

It follows that energy models should be built with the understanding of these dynamics, and the inclusion of these characteristics, so as to improve the quality and reliability of the results and hence provide a more accurate account of the contextual energy situation. However, as argued by Urban *et al.* (2007), the lack of adequate data is a typical constraint to modelling energy demand in developing countries.

Having discussed the main classifications of energy demand models in the literature, the next few sections evaluate the approaches used in relevant literature, under two broad headings: bottom up (engineering or statistical models), and top down (econometric models). As stated earlier, despite the intrinsic merits of each method, both have limitations and drawbacks.

3.4 Bottom up models

The bottom up category is a technology explicit model, where the inputs, outputs, unit costs and other technical and economic characteristics, of each important energy-using technology, are identified and analysed using that information (Loulou *et al.,* 2004). Also, the researchers that have employed this technique have used energy consumption data from a representative household to deduce the estimate at a national or regional level (Swan and Ugursal, 2009). In other words, the bottom up models employ the use of disaggregated data in energy demand

modelling (Urban *et al.,* 2007). This class of models could be classified further into the statistical and engineering approach (Swan and Ugursal, 2009).
In the engineering approach, the demand for a particular fuel type is estimated in relation to the machinery it is used for, how efficient the machinery or equipment is, and the equipment level of usage (Pesaran *et al.,* 1998). Examples of the data used for the modelling include variables such as average energy efficiency, average appliance power ratings, and end use features (Torriti, 2014). Mathematically, as illustrated by Pesaran *et al.* (1998), the demand for fuel (g) using this method, takes the form of the expression below:

$$E_{gt} = \sum_k e_{gk} \varphi_{kt} \, l_{kt} \tag{3.08}$$

In the equation above, l_{kt} is the amount of machinery of type l in use at time t, φ_{kt} is the degree or amount of usage (utilisation) at time t, e_{gk} is the technical energy coefficient of fuel g when used in combination with machinery l. The price and income effect, including the time lags of the variables, in both the short and long term on the demand for energy or fuel is estimated through the analysis of e_{gk}, φ_{kt} and l_{kt}. Examples of this approach used in the residential sector energy consumption modelling include population distribution, archetype and sample (Swan and Ugursal, 2009).

Swan and Ugursal (2009), argue that the engineering method is the only energy model that can analyse a sector energy consumption, without the use of any historical data. They add that the method is especially suited for estimating new technologies with no prior consumption information.

Pesaran *et al.* (1998) point out three main strengths of using this approach. The three things which could be improved through the knowledge gained from the use of the model are: (i) investment in new machineries; (ii) improving the energy efficiency of the existing machineries; and (iii) altering the usage of the existing machineries or equipment. Despite the merits of this approach, it has a major drawback. The approach is data-driven since it requires the use of a considerable amount of data which might not be available especially for developing countries (Pesaran *et al.,* 1998).

Examples of the bottom up model used in the energy literature include the MARKAL, TIMES and the LEAP model. The Long-range Energy Alternative Planning System (LEAP) is used in the study by Ouedraogo (2017) to project energy demand in the four main sub-regions in Africa, for the period between 2010 and 2040. The MARKAL model, stands for 'MARKet ALlocation', is a mathematical model that employs the use of a technology-based technique in energy analysis (Loulou *et al.,* 2004). It was first developed by the Energy Technology Systems Analysis Program (ETSAP), as a least cost linear programming model (Suganthi and Samuel, 2012).

Fishbone and Abilock (1981) provided the equations for the first MARKAL model, which represents both the supply and the demand components of the energy system. The energy system represented by the equations could be estimated at the state, national or regional levels. However, in the last few decades, many improvements have been made to the system of equations so as to facilitate its use for a more in-depth and detailed analysis (Suganthi and Samuel, 2012). Suganthi and Samuel (2012) argue further that the MARKAL model could be used for uncovering the implications of carbon reduction programs and the impact of policy changes (Kanudia and Loulou, 1999, Loulou *et al.,* 2004).

Several researchers have used the MARKAL model to create different scenarios using various policy conditions, to estimate how air pollution emission reduction or the greenhouse gas mitigation are achieved during the implementation of the different policies, and their associated benefits. This was used for Shanghai by Chen *et al.* (2002), Kan *et al.* (2004) and Changhong *et al.* (2006), among others. A more recent study for Taiwan by Tsai and Chang (2013) also took a similar approach, but modelled the GHG mitigation under different technology development scenarios, using the same emission reduction target. The model has also been used for the allocation of energy resources across sectors in an economy (Mallah and Bansal, 2010, in India), and for the analysis of sectoral energy consumption pattern and carbon emissions (see Shrestha and Rajbhandari, 2010, for Nepal).

The statistical approach uses mainly regression techniques for the estimation of end-user energy consumption, like the residential sector, as well as for forecasting different aspects of energy. Swan and Ugursal (2009) grouped the statistical approach into regression, conditional demand analysis and neural network, in the residential sector energy consumption modelling. The authors defined each of the classifications as follows. First, in the regression technique, the coefficients of the input variables in the model are determined through the use of regression analysis (while in the conditional demand analysis the end use appliance data is used as the basis of the regression analysis). The coefficients from the regression analysis of the total residential energy consumption are based on the number of owned appliances, and the result obtained is interpreted as the level of usage of such appliances in the sector. Lastly, the neural network (NN) is developed based on the densely interconnected parallel structure of biological neural networks, which is built into a mathematical model. They argued further that just like in the regression techniques, the model used in the neural network tries to minimise error by accounting for non-linearity through the use of scaling or activation functions (Swan and Ugursal, 2009).

Models used by researchers include multiple regression analysis, panel threshold regression, linear regression models and multivariate regression models. Some of the studies include Tunc *et al.* (2006), who employed the use of regression analysis to forecast Turkey's electricity consumption for the period 2004-2020, using a data set from 1980 to 2001. The electricity consumption of Italy up to 2030, was predicted using the annual historical data of electricity consumption, Gross Domestic Product (GDP), GDP per capita and population in a multiple regression framework, the data used were from 1970 to 2007 by Bianco *et al.* (2009). The least square method was used by Aranda *et al.* (2012) to forecast the yearly energy consumption in the Spanish banking sector, by estimating the regression coefficients using a data set of 55 banks in Spain. Two types of statistical models were used by Kialashaki and Reisel (2013) to forecast the

United States' residential sector energy demand. The multiple regression model and artificial neural network were employed in the energy consumption prediction between 2010 and 2030, using the data from 1984 to 2010.

However, despite the strength of the bottom up modelling technique, its ability to analyse technological options available is limited by the high level of information needed in the model (Swan and Ugursal, 2009). Moreover, the model output could be biased if incomplete technological data are inputted into the system (Urban *et al.,* 2007).

3.5 Top down models

Top down models use econometric modelling techniques, employing aggregated data from economic indicators such as GDP or price elasticities, to estimate and or predict energy demand (Urban *et al.,* 2007). Such models do not provide any information about the state of technology or its level of efficiency but they provide historical references in relation to economic frameworks (Urban *et al.,* 2007). Bhattacharyya and Timilsina (2010) argue that such models are used to validate economic theories empirically, by analysing if there is any significant relationship statistically. The models are used to know how the identified variables impact on the demand for the different energy sources, through the estimation of the independent variables elasticity estimates (Zarnikau, 2003).

However, the approach is most suitable if there is long historical data available on energy consumption, population, income and prices (Pesaran *et al.,* 1998). Pesaran *et al.* (1998) point out that the formulation of an appropriate energy demand equation derived from either households or firms' decision making based on utility optimising behaviour, is used as the theoretical foundation of such analysis. Furthermore, Karimu and Brannlund (2013) argued that most of the econometric models used in the study of energy demand in the literature, range from simple static models to dynamic models, which mostly take a linear or log-linear form. The authors grouped the models into parametric and non-parametric energy demand models. The parametric category is further divided into three forms, which are: log-linear, linear and trans-log models. Karimu and Brannlund (2013) also assert that the log-linear specification is the most commonly used in the analysis by researchers, perhaps due to the ease in its specification and estimation. In this model, before the regression is carried out, both the dependent variable and the explanatory variables are first converted into their natural logarithm forms (Zarnikau, 2003). This will be the approach used in the analysis of the present study, also to ensure that the estimated coefficients can be interpreted as elasticities.

On the other hand, the non-parametric model is a data-driven approach, where the true relationship between the variables is assumed to be provided through a kernel regression (Zarnikau, 2003). This is shown by the equations below proposed by Zarnikau (2003) to illustrate the three functional forms of the parametric models. This equation could be considered by a researcher analysing, for instance, the household demand for electricity in a specific country.

Linear

$$KWH_y = a_x + b_x * PE_y + b_j * PN_y + b_y * INC_y + b_{HD} * HD_y \quad (3.09)$$

Log-Linear:

$$\text{Log } (KWH)_y = a_x + b_x * \log(PE_y) + b_j * \log(PN_y) + b_y * \log(INC_y) + b_{HD} * \log(HD_y) \quad (3.10)$$

Trans log

$$STE_y = a_x + b_x * PE_y + b_j * PN_y + b_y * INC_y + b_{HD} * HD_y \quad (3.11)$$

where KWH_y represents the household's total consumption of electricity, PE_y is the price paid by all households for electricity, STE_y is the total expenditure of household (y) on electricity, HD_y denotes the degree of days of heating, INC_y is the income of the households, PN_y stands for the price of natural gas, which is another source of energy (substitute) for heating purposes.

It appears that some methods are more suitable for aggregate analysis, while others may be more appropriate for a sector or sub-sector specific analysis of energy demand. For instance, Bhattacharyya and Timilsina (2009) point out that three main energy models are used for the analysis of the demand in the transport sector. The three models are: identity models, structural models and the market-share model. However, despite the numerous groupings of energy models, there are some similarities and overlaps among the different models, especially in the hybrid models which involve the use of features from multiple models in the analysis of energy demand. For example, a hybrid model could use both the top down and the bottom up approach in the analysis of residential electricity consumption, and such models have been proven to provide a more detailed level of information (Swan and Ugursal, 2009).

The models are used for both the aggregate and the sectoral level energy demand analysis. The main sectors studied are residential, industrial, transportation and commercial. Some of the studies that used this technique (the econometric approach) for either the aggregate or sector level analysis, are discussed in the next few sections.

3.6 Aggregate demand for energy

The aggregate demand for energy refers to the total energy demanded for all services, which is needed to operate in all economic sectors of an economy. Endresen (2004) argues that the aggregate demand for secondary or converted energy in a country refers to the sum of all energy required in all economic sectors, to make available the energy required for all services that need energy. Endresen (2004, p.104) argued further that

"*a limited source of fuel type could be used to provide some service like commercial road transport, while service such as electricity could be provided using many types of fuel*".

The researcher also asserted that some of the fuel types like petrol could be used directly while others like uranium, in the form of nuclear energy, need to be converted first to a state that can be used for providing energy services. However, some fuels like diesel have a dual use in the sense that they could be used directly for the running of machineries or used indirectly by first converting them into electricity, which is then, in turn, used to run the machineries.

Against this background, some researchers (e.g., Chakravarty, 2002) have attempted to analyse the role played by different factors or variables in total energy demand, the changes in the variables over a period of time under different situations and assumptions, within a parametric or non-parametric energy demand framework. Specifically, some of such studies in the literature include those by Al-faris (1992), Eltony and Hoque (1996), Mohammad and Eltony (1996), Masih and Masih (1996a and b), Brenton (1997), Diabi (1998), Ghali (1998), Pesaran *et al.* (1998), Sinton and Fridley (2000), De Vita *et al.* (2006), and Wolde- Rufael (2006).

Some of the exiting studies in the literature, mostly those in developing countries, will be discussed in more detail later in this chapter. They are based on the law of demand and the assumption that the demand for energy mostly behaves as a normal good, which suggests that the main variables that influence the demand for energy are price and income. This theory has been discussed in the earlier section under the theoretical background of the study (section 3.2).

Furthermore, as pointed out by Samuel *et al.* (2013), the key determining variables for the demand for energy include per capital real GDP, industrial growth, real price of energy, population, air temperature, financial development variables, capital stock, foreign direct investment and efficiency variables. The use of some of these variables, with the price and income variables in the model specification is largely dependent on data availability, and the technique used which has led to different results in terms of size and sign of the elasticities in both the short- and the long-run.

In their study on Namibia, De Vita *et al.* (2006) estimated the demand function at the aggregate level and by specific fuel types (electricity, petrol and diesel), within an autoregressive distributed lag (ARDL) bounds testing approach to cointegration (3.16) derived from equation in (3.15) showing the long run energy demand function for Namibia.

$$\mathrm{e}d_t = \alpha + \beta_1 y_t + \beta_2 p_t + \beta_3 x_t + \mu_t \qquad (3.12)$$

where $\mathrm{e}d_t$ represents the consumpton of energy, y_t stands for GDP, and p_t is the energy price. Quarterly end user data between for the period 1980-2002 were used for the analysis.

(3.16) is used to test the cointegrating relationship among the variables, within an ARDL framework, which is particularly useful in overcoming the problems of mixed order of integration, that is I(0) or I (1), among the regressors (De Vita *et al.,* 2006).

$$\Delta ed_t = c_0 + c_1 t + \sum_{i=1}^{m} \propto_i \Delta x_{t-1} + \sum_{j=0}^{n} \beta_j \Delta y_{t-j} + \sum_{k=0}^{p} \delta_k \Delta p_{t-k} + \sum_{r=0}^{q} \emptyset_r \Delta x_{t-r} + \emptyset D_t + \pi_1 ed_{t-1} + \pi_2 y_{t-1} + \pi_3 p_{t-1} + \pi_4 x_{t-1} + \varepsilon_t \qquad (3.13)$$

where c_0 *and* c_1 represent the intercept and time trend components respectively, and Δ denotes the first difference of each variable. The results from the estimation of the model conform to theory prediction, that is, the price elasticity was negative while the income elasticity was positive, as we would expect *a priori*. The empirical study also looked at the price elasticity of each of the different fuel types, and found that the price elasticity of petrol was the highest, followed by that of electricity, with diesel showing no significant price elasticity. Further, the different energy forms gave no evidence of cross price elasticities in the analysis for Namibia, during the study period. This could suggest that consumers do not necessarily change their consumption level or energy mix according to changes in income or prices of the different energy types, as they appear to maintain the use of certain appliances and equipment for energy generation (De Vita *et al.,* 2006).

Despite the rigorous empirical analysis carried out, no attempt was made in the study at estimating the short run elasticities of the variables, as the study concentrated on the analysis of the long-run. This was the first study in the energy literature to use the ARDL bounds testing approach to cointegration and end-user data for the Namibian economy.

In the analysis of aggregate energy demand, some authors only estimated specific fuel or energy types, but still following the standard econometric approach that includes establishing the variables' integration (unit root) and cointegration properties in addition to the coefficient elasticities. For example, following De Vita *et al.* (2006), Akinboade *et al.* (2008) estimated the demand for gasoline in South Africa between 1978 and 2005, using annual time series data, within an ARDL framework. The results of the analysis show that the demand for gasoline is a normal good. However, the demand increases as the income level increases, but not at a proportional rate.

The same modelling technique of cointegration was employed by Amusa *et al.* (2009) in the analysis of aggregate electricity consumption in South Africa, over a sample period from 1960 to 2007. In contrast to the determinants of gasoline demand by Akinboade *et al.* (2008), the authors found that income is the only factor that drives electricity demand in South Africa, while the price of electricity has no impact during the study period.

Specifically, holding all other factors constant, from the long run income elasticity result of electricity demand, a one percent increase in income leads to a 1.67 percent increase in electricity demanded at the aggregate level in South Africa. Amusa *et al.* (2009) claim that the result of the income elasticity could be linked to electricity consumption and income in three possible ways. First, an increase in the size of the South African economy (GDP) will increase the use of electricity because of its important role in the industrial and manufacturing processes. Second, more durable and electricity driven machinery will be acquired when there is an increase in the production process. Lastly, households will increase the amount of electricity using appliances and services as their income level increases. However, the result of the price elasticity of electricity demand contradicts a prior expectations and the prediction of economic theory, through a positive price elasticity result of 0.30, which is not statistically significant.

The researchers explained these results by using the South African context. They highlighted the lack of no close substitutes, the long period of subsidy by the government, and price reduction strategies used over past decades, as major causes why changes in the price charged for electricity appeared to have no adverse effect on electricity demand in South Africa.

A similar cointegration approach was used for the estimation of petroleum products demand elasticities in Nigeria between 1977 and 2006 by Iwayemi *et al.* (2010). The authors found that both income and price elasticities of petroleum demand in Nigeria had the signs predicted by theory, except those of the diesel demand models. Factors such as inadequate supply that sometimes creates artificial scarcity of the product was used to explain the unexpected elasticity (Table 3.1) results for diesel by the authors. Furthermore, income and price are the two most important factors determining energy demand in the country. Among all the products having an inelastic demand, gasoline recorded the highest income elasticity, followed by kerosene. One possible implication of these results is that the Nigerian government could use tax to increase fiscal revenues and also use this to achieve energy consumption conservation plans (Iwayemi *et al.,* 2010). In other words, tax could be used to control the level and structure of energy consumption in Nigeria.

These results are in agreement with Abdullahi's (2014) findings which also suggest that the demand for petroleum products in Nigeria is both income and price inelastic, and the low elasticities figures for the products could give room for more fiscal revenue generation by the government. Furthermore, in Abdullahi's (2014) study he was able to take into account the impact of structural or technical changes in the analysis through the use of a structural time series model, which was not taken into account in the earlier study by Iwayemi *et al.* (2010). These extensions made the findings more interesting and robust, as it was found that none of the models used for the fuels had a deterministic linear trend.

An earlier study by Expo *et al.* (2011) investigated the dynamics of electricity demand and consumption in Nigeria, also using the same approach to cointegration. The most interesting finding in the study was that price of electricity has no significant impact, as the coefficient was not statistically significant. It seems possible that this result may be due to the involvement of the Nigerian government through electricity price regulation (Expo *et al.,* 2011). Furthermore, GDP per capita, population and industry output were found to be the main factors that explain electricity consumption both in the short- and long-run. The period covered by the study was between 1970 and 2008.

Adom *et al.* (2012) analysed the driving forces of the domestic demand for electricity both in the short- and long-run in Ghana, for the period from 1975 to 2005, within an ARDL framework. The researchers used a log linear model with the use of annual time series data on real capita GDP, industry efficiency, structural changes in the economy, total domestic electricity consumption and the degree of urbanisation variables. Their study does not take account of the effect of the price variables on electricity demand in Ghana during the period of the study due to lack of relevant data. Therefore, the study did not detect any evidence for the impact of changes in price on electricity in Ghana in the analysis. This would have perhaps made the findings more robust.

The inclusion of the impact of structural changes in the analysis is due to the shift towards more energy intensive sectors, which may have led to an increase in the amount of electricity consumed (Adom *et al.,* 2012). The results of the analysis show that, during the study period, real per capita GDP, industry efficiency, the degree of urbanisation and structural changes are the main long-run factors that influence the demand for electricity in Ghana. Surprisingly, the authors found the coefficient of industry efficiency to be negative and significant. The results suggest that energy in the form of electricity is saved, as firms make use of more energy efficient technology in their production process. Thus, reducing the overall energy intensity through a reduction in the industry electricity intensity. In other words, there should be appropriate electricity efficiency policies and regulations for each sector of the country.

These results are consistent with those obtained by Adom (2013) who also found that income, industry efficiency and economic structure are the main factors that impact electricity consumption in Ghana. However, there are some important differences between the two studies. First, the elasticities figures obtained are not the same for the variables, which may be due to differences in the econometric technique employed and the different study period covered by the two studies. Second, Adom (2013) used the Quandt-Andrews test to explore the presence of structural breaks in the data during the study period, and found the year 1987 to be the most likely point of break during the study period. Lastly, he gave evidence of changing policy regimes in his analysis, which made the findings more robust.

The importance of taking into account the impact of policy regime changes was explored further by Adom and Bekoe (2013), in modelling the demand for electricity in Ghana. The authors assert that the inclusion of policy regime changes in the analysis is due to the fact that the decision rule changes because new individuals change their behaviour as new structural policies are put in place (see also Inglesi- Lotz, 2011). The sample used for the analysis was divided into the pre-reform, post-reform and the full-period. The results from the fully modified ordinary least squares regression gave evidence of the pre-reform period being more energy saving than the post-reform period. Furthermore, there seems to be a shift towards the less energy intensive sector after the economic reform in Ghana during the study period between 1971 and 2008.

The estimation of the demand for natural gas in the energy literature could be attributed to its increased importance in the energy mix, and as argued by Ackah (2014), it is cheaper than petroleum and cleaner than coal to burn. This and other factors led to the investigation of the determinants of natural gas demand in Ghana by the above-mentioned author between 1989 and 2009 for the aggregate, residential and industrial sectors. The results of the three models specified by the author using four lags, reveal that all the price and income elasticities had the sign predicted by economic theory.

The debate was taken a step further by Mensah (2014) who explored the modeling of demand for LPG in Ghana using two techniques. Specifically, Mensah (2014) stated that the use of the ARDL and PAM (Partial Adjustment Model) techniques in his study serves the purpose of identifying the best model to be used for a 10 year forecast of the fuel. The quarterly time series data used in both models, between 1992 and 2012, show that income, price and urbanisation are the main determinants of natural gas demand in Ghana in the long run. The author added that the ARDL is a better model for forecasting future LPG consumption. This led to the use of this model by the author to forecast 10 year ex-ante demand for LPG in Ghana, based on the three different scenarios presented by the researcher. The result of the projections from the three scenarios gave evidence that the demand for LPG might, by the year 2022, reach a minimum value of 5.0 million metric tons (Mensah, 2014).

A cross-country study of the factors that drive energy demand in SSA is provided by Keho (2016), who used the bounds testing approach to cointegration at individual country level between 1970 and 2011. The author found that in the 12 countries analysed, economic growth, industrial output and population are the major drivers of energy demand.

Most of the studies reviewed so far in this section of the chapter explored the factors that drive the demand of non-renewable energy sources like crude oil and LPG, and not of the renewable energy sources. This motivated the study by some researchers including Ackah and Renatas (2015), who investigated the determinants of the demand for renewable energy in oil producing economies in Africa. Their findings using three panel data models of random effect, fixed effect and a dynamic model, suggest that the main factors that influence demand are real income, energy resource depletion, carbon emissions and energy prices. The price and the income variables had the expected inverse relationship with demand, as expected theoretically. The main implication of the study for the oil producing countries considered, is the need for the governments in these countries to reduce the barriers in terms of technology and investment climate, so as to attract investors and, in turn, increase the consumption of renewable energy by consumers.

3.7 Disaggregate demand for energy

Recently, researchers have shown a growing interest in the determinants of energy demand of the different fuel types, in different sectors of the economy, in both the developed and developing countries. However, most studies focus on the developed countries. This is also referred to as the sectoral analysis of the energy forms in different sectors of the economy. In the paragraphs that follow, some of the main studies on the energy demand in the residential, industrial and transport sectors in different countries are discussed.

Ziramba (2008) asserts that the most common variables used for the modelling of residential sector electricity consumption are income, price of electricity, price of substitute energy source and temperature variables. However, due to the lack of relevant data for the analysis by Ziramba (2008), he explored the determinants of electricity demand in South Africa using per capita income and the real price of electricity as the sole explanatory variables. This, perhaps, limited the scope and findings of the study, as more relevant variables would have made the results more robust, and may have given more insight into the factors influencing electricity demand. The author used a double logarithmic form model (3.19) for the analysis, following the work undertaken by Narayan and Smyth (2005), who identified income, temperature and price as the determinants of electricity in Australia, for the study period 1969-2000.

$$lnEC_t = \alpha_0 + \alpha_1 t + \alpha_2 lnY_t + \alpha_3 lnP_t + \varepsilon_t \quad (3.\ 14)$$

where lnEC, which is the dependent variable, stands for the natural log of per capita residential electricity consumption (kWh per capita), ln Y is the natural log of real per capita income, lnP represents the natural log of the real residential electricity price (R/kWh), t is the time trend and ε is the error term. Economic theory predicts α_2 and α_3 to be positive and negative respectively, as explained earlier, under the theoretical background.

The bounds testing approach to cointegration within an autoregressive-distributed model was used for the empirical analysis using the data from 1978 to 2005. The author found that the demand for electricity is a normal good, and the price of electricity is price inelastic in both the short- and long-run during the study period. However, due to the statistical insignificance of the price variable, the researcher found that income is the main driver of electricity in South Africa.

Most of the studies reviewed so far used either the linear or the log-linear approach for the estimation of energy demand. However, Lescaroux (2011) modelled energy demand both at the aggregate and sectoral levels for a panel of 101 countries, using a non-linear technique of iterative least squares. The author found that income and price had a significant effect, as predicted by economic theory on the demand for energy, even when a non-linear approach is employed. The impact of changes in income on energy demand was found to be positive, whereas the effect of changes in price is negative on the demand for energy. The author argues further that the impact of changes in income is mostly felt in the industry and service sector, while the residential and the service sectors saturation levels are more affected by changes in the price level.

Evidence of the long-run variables that explain the residential demand for electricity was given by Dramani and Tewari (2014) for Ghana. The authors used a dummy variable in the ARDL model, to account for structural breaks in the demand for electricity. The equation used within an ARDL framework is:

$$lnEC_t = \propto_0 + \sum_{t=0} m \propto_1 lnRPE_{t-i} + \sum_{i=0} n \propto_2 lnPC_{t-i} + \sum_{i=0} p \propto_3 lnUP_{t-i} + \sum_{i=0} q \propto_4 lnPK_{t-i} + \sum_{i=0} r \propto_5 lnLPG_{t-i} + \sum_{i=0} s \propto_6 lnEI_{t-i} + \theta DU_{t-i} + \in_t \tag{3.15}$$

where EC is electricity consumption, RPE stands for the price of electricity, LPG represents the price of gas, PC is per capita GDP, UP stands for urbanisation, EI is the intensity of the residential consumption of electricity, PK denotes the price of kerosene, $\in_t$ is the error term. The price of gas and kerosene is included as alternative energy sources to electricity, and the result gave evidence that electricity demand in Ghana is inelastic. In other words, consumers in the residential sector consider electricity as a necessary commodity and kerosene and gas would only be used as a complementary of electricity by them (Dramani and Tewari, 2014).

All the variables were found to be cointegrated and the results of the ARDL model showed that the variables have all the expected signs. The coefficient of the degree of urbanisation was found to be positive and significant statistically, which could be because urban dwellers could acquire more appliances because of their increased chance of securing a higher paying employment than the people who live in the rural areas- this is the case for most of the countries in Sub-Saharan Africa.

Dramani and Tewari (2004) argued further that due to a more improved electricity markets and distribution systems, the positive impact of the effect of increased urban population on electricity consumption, could be linked to better access to electricity and appliances utilisation.

The study was also interesting due to the use of FMOLS and DOLS techniques employed by the researcher to correct for endogeneity bias. However, due to some differences in the results of the three models, and also the statistical insignificance of some of the variables in the FMOLS and DOLS, the researcher argued that the ARDL is more robust (see also Mensah, 2014) than the two other models, because it is not dependent upon the econometric technique used, and - therefore - provides the best framework for estimating Ghana's electricity demand function.

The determinants of household use of clean and renewable energy sources for lighting in SSA were investigated by Rahut et al. (2017), using Ethiopia, Tanzania and Malawi as case studies. In the multinomial logit and ordered probit models employed by the authors, the results suggested that wealthier and more educated households use electricity and solar energy for lighting, whereas, poorer consumers use solid fuels, batteries and kerosene for lighting. They also found that households with female as the head are more likely to use clean and renewable sources of energy than those with male head, in the countries analysed.

3.8 Energy demand studies in non SSA countries

Kumar (2008) investigated the determinants of energy demand in Fiji for the period 1970-2005. Using a four period lag length, he found a cointegrating relationship between energy prices, GDP and energy consumption. Both income and energy prices had the expected elasticities signs and were both statistically significant. The unit income elasticity of energy demand suggests that energy demand rises with the same proportion with the change in the income level. The time series techniques within a General to Specific (GETS) and Johansen Maximum Likelihood (JML) were employed for the analysis by the author. The JML was further used by the author to test for the direction of causality among the variables both in the short- and in the long-run. The author found that both in the short and in the long run economic growth led to energy consumption. In other words, there is a unidirectional causality running from GDP to energy consumption in Fiji, both in the short- and long-run.

In contrast to the findings in other countries in the developing region, Alter and Syed (2011) found no evidence that the demand for electricity is a normal good both in the short and long-run, in Pakistan. The authors found that over the 1970-2010 sample period, the demand for electricity behaves as a luxury good in the long run. From their analysis using the Johansen cointegration approach, the authors found that as regards to the price elasticities, for all the sectors analysed, only the aggregate and industrial analysis result confirmed that the demand for electricity is a necessity. Whereas, in the residential, commercial and the agricultural sectors, it appeared to be a normal good. A similar observation was also recorded in the income elasticities, where only the aggregate and the commercial sectors showed that the demand for electricity in these sectors in Pakistan is a necessity. These results may be explained by the fact that only 30 to 40% of the entire population in Pakistan is connected to the national grid, as confirmed by the studies by Khan and Qayyum (2008) and Jamil and Ahmad (2010).

Zachariadis and Pashourtidou (2007) investigated the electricity use in the residential and the service sectors in Cyprus, using a data set between 1960 and 2004. Income, prices and weather were used as the regressors in the analysis. The inclusion of the weather variable, proxied by the degree days, is due to the use of air conditioning during the high temperature seasons, and likewise the use of energy for heating when the temperature becomes very low. The VEC equation used is:

$$\Delta e_t = \propto_{01} + \propto_{11} \Delta e_{t-1} + \propto_{21} \Delta y_{t-1} + \propto_{31} \Delta P_{t-1} + \propto_{41} \Delta tdd_t + \propto_{51} (e_{t-1} + by_{t-1} + cp_{t-1} + d) + u_{1t} \quad (3.16)$$

where e stands for electricity consumption, p is the price variable, y represents income, and tdd is the total degree-days variable. The error correction term derived from the cointegration analysis is in parenthesis.

The researchers found that electricity in Cyprus has no close substitutes, as all the tests for competing fuels were insignificant. Furthermore, the inclusion of 'degree-days' as the weather variable seems appropriate, because of the statistical significance of the coefficient. The researchers also analysed the direction of causality among the factors that influence electricity consumption. The results show that in the short- and long-run, in the residential sector, there is a bidirectional causality running between private income and residential electricity consumption. However, no Granger causality was observed in the variables in the model for commercial electricity consumption.

Al-Azzam and Hawdon (1999) analysed energy demand in Jordan using a linear log model in the estimation of the Jordan energy demand with a dataset based on the 1968-1997 sample period. The Stock-Watson Dynamic Ordinary Least Squares (DOLS) and the Error Correction Model (ECM) techniques were used to investigate the relationship between energy consumption, real income, real energy prices and construction activity. The long-run specification was:

$$\text{Log}\,(Q_t) = 0 + 1\log(P_t) + 2\log(Y_t) + 3\log(A_t) + 4D_i + t \qquad (3.17)$$

where Q_t stands for the aggregate energy consumption, P_t represents the real price of energy, Y_t is the real income (GDP), A_t stands for the total area constructed in square meters. D_i stands for a dummy variable for conflict and political stability, 0 is the intercept, 1,2 ,3 and 4 are the elasticities of total energy demand, and t is the error term.

The authors found that the best method to obtain a robust result was the DOLS technique due to its ability to work well even in the presence of a mixed order of integration in the variables, small sample size, and also to correct the problems found when using the error correction model framework, hence giving more reliable estimates of the elasticities. Further, both the Johansen Maximum Likelihood (JML) and the DOLS techniques confirm that the variables were co-integrated in the long run. The price elasticities of energy demand were found to be low and statistically insignificant. The income elasticity of the total energy demanded was found to be close to unity, suggesting that the amount of increase in the demand for energy induced by the growth in the economy is proportional.

There are a number of similarities between the studies by Al-Azzam and Hawdon (1999) and Saed (2004), as both investigated the factors influencing energy demand in Jordan using the same techniques (DOLS and ECM), and they also included dummy variables in their models to account for conflict and political instability during the study period. Moreover, the results of both studies in terms of the income elasticities of energy demand show that there is a positive relationship between energy consumption and income.

However, there are a number of important differences between the two studies, including the data sets used, the size of the long run price elasticities obtained, and also the implications. While Al-Azzam and Hawdon (1999) found that energy demand is not price responsive, Saed (2004) found that energy demand is price responsive and conservation policies could be achieved through the use of taxation. This inconsistency could be as a result of the different datasets used for the analysis. Furthermore, another major drawback of the approach used by Al-Azzam and Hawdon (1999), and Saed (2004), is the failure to estimate the short run linkages among the variables in the study. Perhaps, this would have made the findings more robust and interesting in terms of being comparable to other studies.

Alves and Bueno (2003) studied the short run, long run and cross elasticities of gasoline demand in Brazil, using price, income and the price of alcohol as independent variables, within a co-integration model. Alcohol-based fuel is chosen because in Brazil it has been developed as a close substitute to gasoline, and it is commonly used as a major fuel for automobiles. The study covered the period between 1974 and 1999 for all the variables, including gasoline consumption per capita, real per capita GDP, and the real price of gasoline. They found that both price and income had the expected sign, as predicted by demand theory. Also, a positive sign for the cross-price elasticity of gasoline for alcohol was recorded. The result of the cross-price elasticity of alcohol and gasoline from the error correction model, may suggest that in the long run consumers are not very sensitive to changes in the price of fuel. This is in line with the hypothesis that the price elasticity of gasoline is inelastic in both the short- and long-run.

Alves and Bueno (2003) estimated the cointegrating equation:

$$\ln C_t = \beta_0 + \beta_1 lnY_t + \beta_2 lnP_t + \beta_3 lnA_t + e_t \qquad (3.18)$$

where C_t stands for the yearly gasoline consumption per capita measured in liters, Y_t represents the yearly real per capita GDP, P_t is the annual real gasoline price, A_t is the yearly real alcohol price while the e_t is the error term.

(3.13) long run elasticities are given by:

$$\frac{\partial lnC_t}{\partial lnY_t} = \beta_1; \frac{\partial lnC_t}{\partial lnP_t} = \beta_2; \frac{\partial lnC_t}{\partial lnA_t} = \beta_3 \quad (3.19)$$

β_1, β_2 and β_3 represent the income elasticity, gasoline price elasticity and the alcohol cross-price elasticity, respectively.

Dilaver and Hunt (2011) forecasted the demand of electricity in the residential sector in Turkey, using a structural time series model between 1980 and 2008. The model used was:

$$E_t = f(Y_t, P_t, UEDT_t) \quad (3.20)$$

where E_t represents the residential electricity demand, Y_t stands for the real household total final expenditure, P_t is the real residential electricity price and UEDT is the residential sector underlying energy demand trend.

The authors concluded that household total final consumption expenditure, real energy prices alongside underlying energy demand trends, are the main factors that determine the demand for energy.

Sharma *et al.* (2002) studied the demand for three main forms of energy using a log-linear form, in four main sub-sectors (residential, services, industries and miscellaneous) in the state of Kerala in India. An OLS regression was employed, with the residential sector model taking the functional form:

$$\ln(ER_i/P_i) = \beta_0 + \beta_1 \ln(Q_i/P_i)/(Q_B/P_B) + \beta_2 \ln(N_i/N_B) \quad (3.21)$$

where ER is the energy demand, P is the population, Q is the state domestic product, and N represents the number of electrified residential dwellings.

Similar models for each sectors were specified but with different explanatory variables used as appropriate for each sector under study. The coefficient of determination (R-squared) of the models, showed that the models provided a 'good fit'. An important finding was that by the year 2020, the energy consumption in Kerala will be four times the demand recorded in 2002, when the study was carried out.

Hung and Huang (2015) studied the dynamic demand for electricity in Taiwan under seasonality and increasing block pricing. A significant difference in the electricity demand was found between the summer and non- summer periods. This result highlighted the importance of the inclusion of a climate variable in the analysis. Also, a larger amount of electricity was consumed in the summer months than in the non-summer months. A possible explanation for this observed pattern, according to the researchers, is the need for cooling by the use of air conditioners in the summer months, and also because it might be difficult for consumers to reduce their electricity consumption level when the weather is hot. This suggests that the higher the temperature, the higher the amount of electricity consumed - a situation rather similar to that of Sub-Saharan African countries with more periods of hot weather. The results from the study also gave evidence of increased use of electricity due to the higher price elasticities of the non-summer months from the analysis.

Having reviewed critically some of the literature on both the aggregate and the sector level energy demand, it is important to discuss the direction of the causality between energy consumption and economic growth in an economy. And also, how energy efficiency is important in the debate and, more importantly, the factors that influence its impact on energy consumption. These two important aspects, and the studies that have analysed them, will be discussed in the subsequent sections below.

Table 3.1 summarises some of the elasticities reported by the studies discussed in this chapter, while more work has been done in Asia and the Middle East on energy demand analysis. The reviewed studies in Sub-Saharan Africa show that most of the work in the region focused on individual countries, where either a particular fuel type or sector was estimated. However, the findings from individual countries could be used for the understanding of the demand function for energy in all the countries in the SSA region, as they all have similar energy issues that include the high rate of traditional use of energy sources like biomass, a high number of informal energy institutions, and similar energy policies.

Table 3.1: Table showing the empirical results of selected studies on energy demand

Author	Sector/fuel type	Elasticities Price LR	Price SR	Income LR	Income SR	Regressors	Period	Method	Country
Diabi (1998)	Electricity	0.01		0.09		Price income Urbanisation Appliances price Weather	1980-1992	OLS	Saudi Arabia
Al-Azzam and Hawdon (1999)	Energy	-0.22 0.95		-0.08 0.98		Price Income Total area	1968-1997	JML DOLS	Jordan
Lundmark *et al*. (2001)	Electricity	0.51		0.86		Income Coal prices Electricity price	1980-1996	OLS	Namibia
Alves and Bueno (2003)	Gasoline	-0.09		0.12		Income Gasoline price Alcohol price	1974-1999	ECM	Brazil
De Vita *et al.* (2006)	Energy Electricity Petrol Diesel	-0.34 -0.30 -0.86 -0.11		1.27 0.59 1.08 2.08		Price Income Air temperature HIV/AIDS rate	1980-2005	ARDL	Namibia
Atakhanova and Howie (2007)	Aggregate Residential Service Industrial	-0.04 -0.21 -0.12 -0.07		0.37 0.12 0.76 0.41		Price Income Population Structural changes Efficiency improvement	1994-2003	GMM	Kazakhstan
Zachariadis and Pashourtidou (2007)	Residential Commercial	-0.30 -0.43		1.12 1.11		Price Income Weather	1960-2004	VECM Granger Causality	Cyprus
Akinboade et al.(2008)	Gasoline	-0.47		0.36		Income Gasoline price	1978-2005	ARDL	South Africa

Author	Sector/fuel type	Elasticities Price LR	 Price SR	 Income LR	 Income SR	Regressors	Period	Method	Country
Kumar (2008)	Energy	-0.30		1.00		Price Income	1970-2005	GETS and JML	Fiji
Amusa *et al.* (2009)	Electricity	-11.41		1.67		Price Income	1960-2007	ARDL	South Africa
Iwayemi *et al.* (2010)	Energy Gasoline Diesel Kerosene	-0.11 -0.06 0.11 -0.12		0.66 0.75 0.11 -0.12		Income Price	1977-2006	Cointegration	Nigeria
Alter and Syed (2011)	Aggregate Residential Industrial Commercial Agricultural	-0.19 -0.42 -0.21 -0.30 -0.14		0.32 0.18 0.06 0.00 0.72		Income Price Electric appliance stock Number of customers	1970-2010	Cointegration	Pakistan
Adom *et al*. (2012)	Electricity				1.59	Income Industry efficiency Structural changes Urbanisation	1975-2005	ARDL	Ghana
Adom (2013)	Electricity			2.12		Income Economic structure Industry electricity intensity	1971-2008	Phillip-Hansen	Ghana
Adom and Bekoe (2013)	Electricity			0.81		Income Industry output Industry energy efficiency Urbanisation	1971-2008	Phillip-Hansen	Ghana

Author	Sector/fuel type	Elasticities Price LR	 Price SR	 Income LR	 Income SR	Regressors	Period	Method	Country
Abdullahi (2014)	Gasoline Diesel Kerosene Fuel oil LPG	-0.23 -0.30 -0.20 -0.18 -0.58		0.11 0.17 0.10 0.27 0.64		Price Income	1978-2010	STSM	Nigeria
Ackah (2014)	Gas aggregate Residential Industrial	-1.81 -0.52 -6.3	-0.36 -0.47 -0.41	1.95 0.53 3.70	0.38 0.48 0.25	Price Income population	1989-2009	STSM	Ghana
Mensah (2014)	LPG	-0.28		0.45		Price Income Urbanisation	1992-2012	ARDL PAM	Ghana
Dramani and Tewari (2014)	Residential Electricity	-0.08 -0.15 -0.14	- 0.05	0.94 0.44 1.38	0.60	Price Income Urbanisation Intensity of residential consumption	1970-2010	ARDL FMOLS DOLS	Ghana
Mensah *et al* (2016)	Gasoline Diesel LPG Kerosene Biomass Residual fuel oil Electricity	-0.55 0.32 -0.26 -0.48 -0.76		1.32 3.56 2.77 -3.63 -0.59 1.74 2.71		Price Income Urbanisation Economic structure	1979-2013	ARDL	Ghana

3.9 Energy consumption, economic growth and carbon emission

Existing research recognises the critical role played by energy consumption in the economic growth of countries, in both developed and developing countries, starting from the seminal paper by Kraft and Kraft (1978) on the nexus between these variables for the USA. In the last few decades, this has been the focus of some researchers in the energy economics literature, especially in Sub-Saharan Africa with a very low energy consumption rate, which may be measured by the low electrification rates in the region. Furthermore, the growing level of income in these countries has led to an increase in electricity consumption (Ouedraogo, 2010), which could also explain the rise in the amount of research in this area. Lastly, the energy dependence of some of the SSA countries, resulting perhaps from the increase in economic activity in the region, reinforces the importance of a reliable long run demand trend estimate to plan the needed energy supply (Odhiambo, 2009); thus highlighting the importance of the aim of this research project.

Ozturk (2010) categorised the direction of the causality between energy consumption and economic growth in a survey of the literature on energy consumption and GDP growth between 1978 and 2009 into four main categories, based on their implications in the formulation of appropriate energy policies.

The first category is referred to as the *'neutrality hypothesis',* according to which there exists no linkage between energy consumption and economic growth. That is, the use or consumption of energy in an economy will have no impact on the growth of the economy. Examples of studies that support the neutrality hypothesis in the energy literature for the US, include the work of Akarca and Long (1980) for the period between 1950 and 1970, using the Sim's technique. Yu and Jin (1992), on the other hand, used the cointegration and the Granger causality test for datasets between 1974 and 1990. Payne (2009) in an earlier study for the US used the Toda- Yamamoto causality test for the period 1949-2006. Soytas and Sari (2009) gave the same evidence for Turkey also using the Toda-Yamamoto causality test, between 1960 and 2000. Evidence for SSA countries was added to the literature in the study by Menegaki and Tugcu (2016), for the period between

1985 and 2013. Furthermore, evidence for the neutrality hypothesis was provided in the study between 1971-2012 for South Africa, using GMM by Maladoh Bah and Azam (2017).

The second category pertains to studies which found that the conservation or reduction in energy use will have no negative effect but a positive impact on the growth of GDP. This is referred to as the '*conservation hypothesis'*. This hypothesis found support in the studies by Kraft and Kraft (1978) for the USA over the period between 1947 and 1974, Aqeel and Butt (2001) for Pakistan using a data set between 1955 and 1996, Ang (2008) for Malaysia using Johansen cointegration and VEC models between 1971 and 1999, and Zhang and Cheng (2009) for China, among others.

Thirdly, is the '*growth hypothesis'*, which postulates that energy could serve as a complement to both labour and capital to enhance economic growth in the production function. In other words, an increase in the use of energy would lead to an increase in GDP growth and, likewise, a reduction in the amount of energy consumed will reduce the growth rate of the economy. Studies in support of the growth hypothesis include those by Oh and Lee (2004) for Korea, Wolde-Rufael (2004) for Shanghai, Lee and Chang (2005) for Taiwan, Ho and Siu (2007) for Hong Kong and Iyke (2015) for Nigeria.

The last category is the *'feedback hypothesis'*, according to which there is a two-way relationship between energy consumption and economy growth. This would mean that both energy consumption and economic growth affect one another. See for example, Hwang and Gum (1991), Glasure (2002), Paul and Bhattacharya (2004), Erdal *et al.* (2008) and Menegaki and Tugcu (2016).

The verification of these hypotheses and other factors mentioned earlier, motivated the study by several researchers, including Wolde-Rufael (2006) who explored the long run causal relationship between electricity consumption and economic growth in 17 African countries, using a dataset between 1971 and 2001. Using the bounds test approach to cointegration, the author found that a long run relationship exists between electricity consumption and economic growth in nine countries. Specifically, the results suggest that in Benin, Cameroon, Congo Republic, Gabon, Morocco, Nigeria, South Africa, Zambia and Zimbabwe, economic growth and electricity consumption are cointegrated in the long run. As with most of the studies that analyse the relationship between economic growth and energy consumption in the literature, Wolde-Rufael (2006) explored further the direction of the causality between the two variables in the countries he studied. The results show that in Cameroon, Ghana, Nigeria, Senegal, Zambia and Zimbabwe, there is evidence of a one-way causality running from economic growth to electricity consumption.

However, according to the author, caution should be applied in the interpretation of the results, due to the severity of the energy crisis in the countries in question, as the use of energy conservation policy may worsen the energy crisis experienced in those countries. Furthermore, Wolde-Rufael (2006) asserts that the results of his study may be particularly interesting for the case of Nigeria and Senegal with high electric power transmission and distribution losses if appropriate conservation policies are implemented, which might help to improve the energy efficiency levels in these countries.

In Benin, the Democratic Republic of Congo and Tunisia, the evidence from the study suggests that a unidirectional causality runs from electricity consumption to GDP. Whereas in the case of Eygpt, Gabon and Morocco, evidence of a two-way causality was found on the basis of a modified Wald test (Wolde- Rufael, 2006). The conclusions and findings by Wolde-Rufael (2006) would have been more interesting if other variables were included, as it is not only energy consumption that impact on the economic growth in an economy. Perhaps, the inclusion of other complements like capital and labour could have led to a more convincing result (Wolde-Rufael, 2009). Moreover, the study findings might have been affected by the omission of variables bias which happens in bivariate causality models (Odhiambo, 2009; De Vita and Trachanas, 2016).

Chontanawat *et al.* (2006) noted that the lack of consistent results across countries may be due to institutional differences among them. The lack of uniformity in the results for individual country by researchers, is surprising leading to different explanations by authors in the energy literature. According to Masih and Masih (1997), as cited by Chontanawat *et al.* (2006), possible reasons for the differences in findings, for example may be due to differences in the way variables are specified, the econometric approach employed, number of lags used, and other methodological issues.

To exemplify this further, the study by Akinlo (2008) investigated the causal relationship between energy consumption and economic growth in 11 SSA countries, within an ARDL framework and applying Granger causality tests. Akinlo (2008) found evidence of no causality between energy consumption and economic growth in Nigeria, Kenya and Togo. These results contradict those by Wolde-Rufael (2006) who suggested that an increase in economic growth would lead to more electricity consumption in Nigeria. Also, while Wolde- Rufael (2006) found that in Ghana there is a unidirectional relationship from economic growth to electricity consumption, Akinlo (2008) found evidence of a bidirectional relationship between economic growth and energy in Ghana.

However, as suggested earlier, the differences in findings should also be interpreted with caution because of differences in the sample periods, econometric techniques used for the causality tests, and the definition and measure specification of the variables employed for the analysis. The inclusion of two control variables by Akinlo (2008), made the study more robust and interesting. Specifically, government expenditure and price variables were added by the author to prevent the problem of simultaneity bias and perhaps omission of variables bias in bivariate models, which could make the results incorrect when testing for the direction of causality between energy consumption and economic growth (Akinlo, 2008).

The findings in the study by Wolde-Rufael (2006) were revisited by Wolde-Rufael in 2009, and the model used was made more robust by the inclusion of two additional explanatory variables. The same datasets were used as in 2006, but the author added labour and capital as additional variables, and analysed the data in a multivariate framework. The findings led to two interesting points. First, in most of the countries under study, there was a reverse of the direction of causality when compared to the earlier study. Second, the author found that labour and capital are the two most important determinants of economic growth.

According to Wolde-Rufael (2009), four possible factors accounted for the difference observed in the two studies as regards to the direction of causality. First, is the fact that the bivariate case might lead to bias as there are other factors that influence economic growth, and thus the use of just two variables may be inadequate to capture the factors that impact on output growth. Wolde-Rufael (2009) argues further that the consumption of energy in a country, is inadequate to cause economic growth in that country, as suggested in the bivariate case. This further confirms the omission of variables bias associated in the energy literature to the use of bivariate models.

The second argument is that the implementation of economic liberalisation in some of the sample countries such as Nigeria and Sudan, could have resulted in an increase in the demand for energy through an increase in the economic growth recorded in these economies. This view is corroborated by the argument put forward by Wolfram *et al.* (2012), who assert that the recent increase in the economic growth of developing countries, has led to more demand for energy through improvements in their economic condition.

The third factor pointed out by Wolde-Rufael (2009), is the manner and pace in which the privatisation policies were implemented in some of the countries. As regards to this, the author citied Ghana and Zimbabwe as two countries where a gradual move to privatise the power sector has led to higher economic growth, in contrast to Kenya, where the process seems to be have been hurried and has led to more crisis in the availability of energy. Lastly, the wide variation in the level of energy efficiency in the countries was also used to explain the reversed causality in the new study by Wolde-Rufael (2009). However, despite the merits of the study, the use of the cointegration approach would have made the findings in the multivariate model more robust and convincing, especially when compared to the univariate framework used in the earlier study.

Very little is known about the relationship between energy consumption and economic growth at the sub-regional level. This was explored in the study by Kebede *et al.* (2010), who used cross-sectional time series data for 20 countries from 1980 to 2004. The total energy demanded is measured by the addition of wood fuel (traditional energy), electricity and petroleum. The model used is:

$$ED = f\ (GDP, OPR, PGRT, AGR, IND) \tag{3.22}$$

ED represents energy consumption, GDP stands for real GDP, OPR is world crude oil price, PGRT is population growth, AGR represents value added to agriculture, while IND is value added to industry. The three models used for the regression in the analysis are:

$$LED_{it} = \beta_0 + \beta_1 LGDP_{it} + \beta_2 LOPR_{it} + \beta_3 PGRT_{it} + \beta_4 GR_{it} + \beta_5 IND_{it} + \varepsilon_{it} \tag{3.23}$$

$$LED_{it} = \beta_0 + \beta_1 LGDP_{it} + \beta_2 LOPR_{it} + \beta_3 PGRT_{it} + \beta 4AGRit + \beta_5 IND_{it} + \beta_6 DC + \beta_7 DE + \beta_8 DS + \varepsilon_{it} \tag{3.24}$$

$$LED_{it} = \beta_0 + \beta_1 LGDP_{it} + \beta_2 LOPR_{it} + \beta_3 PGRT_{it} + \beta_4 GR_{it} + \beta_5 IND_{it} + \beta_6 DC + \beta_7 DE + \beta_8 DS + \beta_9 DC*PGRT + \beta_{10} DE*PGRT + \beta_{11} DS*PGRT + \varepsilon_{it} \tag{3.25}$$

In the equations above, i and t stand for country index and time, respectively. LED_{it} is the log transformation of energy consumption in country i at time t, $LGDP_{it}$ and $LOPR_{it}$ represent the log of real GDP and crude oil price in country i at time t. A dummy variable is used to distinguish one sub-region from the others in the models used. The results show that energy demand is a normal good and its demand is highly inelastic in all the countries considered. In other words, because the consumers in these countries depend on energy for many activities, a one per cent change in the price of oil will lead to a 0.1% fall in the quantity demanded of oil, as expected theoretically. Furthermore, evidence of a positive relationship between energy demand and the rate of population growth was found in all the sample countries. Likewise, agricultural expansion was found to lead to an increase in energy demand while a negative relationship was found between the industrial share of value added and energy demand.

Evidence for South Africa was given using time series data in the study by Odhiambo (2009), which covered the period between 1971 and 2006. This study employed a trivariate framework that included electricity consumption, real per capita GDP and employment level as explanatory variables, and used a Granger causality test. The researcher found that there exists a bidirectional causality between electricity consumption and economic growth, in both the short and long-run in South Africa. Furthermore, a unidirectional causality from employment to economic growth was also found in the study. However, the findings differ from Menyah and Wolde-Rufael (2010), who found a unidirectional causality running from CO2 emissions to economic growth, and from energy consumption to economic growth. Furthermore, unlike Odhiambo (2009), who used a trivariate model, the study employed the use of a multivariate framework and a cointegration approach for the analysis. The results of Menyah and Wolde-Rufael (2010) for South Africa, are corroborated by the findings in the study by Odhiambo (2010), who explored the causal relations between energy consumption, prices and economic growth in three SSA countries: South Africa, Kenya and the Democratic Republic of Congo. Using the ARDL approach and the Granger causality test, the researcher found that there is a unidirectional causal flow from energy consumption to GDP growth in both South Africa and Kenya, while as in Congo (DRC) it is economic growth that leads to an increase in energy consumption.

In the case of Burkina Faso, Ouedraogo (2010) used the ARDL and a Granger causality test to examine the causal directions between electricity consumption and economic growth, between 1968 and 2003. The author found that both in the short- and long-run, there is a two-way relationship between electricity consumption and real GDP. This would suggest that both energy use and the growth in the economy complement one another (Ouedraogo, 2010).

Evidence for Ghana, Senegal and Morocco was given in the study by Adom *et al.* (2012). However, the authors deviated slightly through the inclusion of carbon dioxide emissions, technical efficiency and industrial structure, and the omission of energy consumption in their specified empirical model. The short- and long-run relationship among the variables was analysed through the use of the ARDL bounds cointegration test, while the causal dynamics was explored using the Granger causality test. For the three countries, there was evidence of a long-run cointegrating relationship among the variables. For Ghana, during the study period, carbon dioxide emissions in the long run could be explained by industrial structure, technical efficiency and economic growth. The findings would have been more interesting if the authors included energy consumption in the analysis, especially considering the rigorous econometric procedure employed.

Researchers have also explored the relationship between economic growth, carbon dioxide emissions and energy consumption within a bivariate and multivariate panel data framework. Ozturk *et al.* (2010) investigated the relationship between energy consumption and economic growth, using a panel data set of 51 low- and middle-income countries between 1971 and 2005. The authors divided the countries into three groups based on their level of income: low-income group, lower-middle-income group, and upper-middle-income group, so as to resolve the 'lump-together' problem associated with the use of panel data. The results of the panel cointegration test show that in all the three income groups, energy consumption and economic growth are cointegrated. The Granger causality test gave evidence of a unidirectional causality running from economic growth to energy consumption in the low-income group, whereas in the two middle-income countries, there exists bidirectional causality between energy consumption and GDP. However, for all the income groups the authors found no evidence to support that energy consumption could cause economic growth. A serious weakness of this study, however, is the use of a bivariate model (see also Narayan and Narayan, 2010; and Ouegraogo, 2013). The inclusion of new variables such as real gross fixed capital formation, labour force, carbon dioxide emissions, GDP deflator, population, exchange rates, interest rates (Ozturk, 2010), within a multivariate model would have made the findings more convincing and interesting.

Other studies, including Mulali and Sab (2012), investigated the impact of energy consumption and carbon emissions on GDP growth and the financial development of a country. The study by Mulali and Sab (2012) covered SSA over the period 1980-2008, in a panel data model context. It is interesting to note that in all the thirty SSA countries in this study, it was found that primary energy consumption had a positive impact on both financial development and GDP growth. The panel cointegration method was used for the analysis, which seems to be much more robust than the time series cointegration method, because of the larger dataset across both time and space.

A multivariate panel data framework was employed in the study by Jebli *et al.* (2015) to analyse the relationship between carbon emission, gdp growth, renewable energy consumption and trade in 42 SSA countries. The findings of the study gave evidence of a bidirectional relationship between economic growth and carbon emission in the short run. Esso and Keho (2016) also investigated the long-run and causal relationships among energy consumption, carbon dioxide (CO2) emissions and economic growth in a sample of 12 SSA countries, between 1971 and 2010. The authors used bounds test to cointegration and Granger causality tests in their analysis. They found that in most of the analysed countries in the long run, energy consumption and economic growth are linked with increase in carbon emission in most countries.

It is worth noting that despite the large amount of work that has been carried out in the energy literature on identifying the true direction of the causality between energy consumption and economic growth, some of which have been reviewed critically in this section of the chapter, the studies have all led to conflicting and mixed results, which suggests that the debate is continuing with no clear consensus on what the right direction of causality is.

3.10 Energy intensity

Some authors took the energy demand and the debate on the causality between energy consumption and economic growth further, by exploring the factors that determine the rate or level of energy intensity in a country, which could be defined as the ratio of energy consumption to total output (Wu, 2012). This definition is also referred to as the aggregate measure of the total energy use in GDP (Allcott and Greenhouse, 2012).

Despite the numerous practical advantages associated with increased access to modern energy, the role played by energy as a prerequisite for the reduction in the poverty level, and for the realisation of the sustainable development goals (SDGs), including higher economic growth and development, especially in the developing countries context, cannot be overemphasised (Ouegraogo, 2013). However, increased access to modern energy also has some adverse effects, the most serious disadvantages seem to be the environmental changes, which are linked to the pollution of the air, water and soil (Wu, 2012). These issues have motivated some researchers, policymakers and other stakeholders to look for ways to reduce such adverse effects, especially the part contributed by the increase in energy consumption, mainly through the increase in the level of energy efficiency. To exemplify this, a report on the GHG emissions in 2010 showed that 35% of the emissions came from the energy sector (IPCC synthesis report, 2014). Energy efficiency is also believed to provide a platform where both stringent energy conservation policies and good ways of reducing negative impacts of energy use can be achieved (Allcott and Greenhouse, 2012*)*.

Based on the foregoing, it appears that the implementation of a good energy efficiency policy in a country has several benefits, and it may also help in meeting the growing demand for energy, through saving some of the available energy (Wang *et al.,* 2012). Most of the studies on energy intensity for a developing region are from China due to the high level of economic growth, and China's position as the largest consumer of energy in the world (Wu, 2012). The authors who have explored this topic have identified several factors that influence the energy intensity level, and have also investigated the impact of the energy efficiency policy pursued. It should be noted that energy efficiency is the inverse of the level of energy intensity, and, therefore, the former could be used as a good indicator of the latter. Nevertheless, following earlier work (e.g., Wu, 2012), many researchers have sought to shed more light on the factors that drive energy intensity in an economy. The model used by Wu (2012) to investigate the driving forces of the energy intensity level in China's regional economies is explained below:

$$Y_{it} = \propto_0 + \mathcal{E}_j \propto_j X_{ijt} + \in_{it} \qquad (3.26)$$

where Y_{it} stands for either the efficiency index or the structural change index for region i at time t, X_{ijt} represents region specific characteristics.

Other variables in the model include price, income, the capita-labour ratio and the growth rate of capital stock. The capita-labour ratio is used to represent the level of technology. The author (Wu, 2012) posits that growth in an economy could bring about an improvement in the level of energy efficiency, which was measured by Wu (2012) using per capita gross regional product (GRP). Price is expected to have a negative coefficient that is, as the price of energy increases the level of energy intensity reduces. In other words, the energy efficiency level will increase as the price paid for energy increases, which may be because consumers are expected to reduce the amount of energy used due to a higher cost, or buy more energy efficient equipment and gadgets.

Using the generalised method of moments (GMM) methodology, the author found a decrease in the average level of energy intensity in China over the study period between 1997 and 2007. This finding suggests that the energy efficiency level has improved during the study period. However, at the regional level there seem to be differences, with respect to the rate at which an increase in the rate of energy efficiency was achieved. In other words, while in some regions there was evidence of a dramatic decline in the energy intensity level, an increase of up to seven times higher than the lowest was found in others. Furthermore, Wu (2012) argued that income, energy prices and the introduction of new technologies are the main factors that determine the energy intensity level in an economy.
Evidence on a disaggregated basis from the industrial sector in China was given in the study by Wang *et al.* (2012), who investigated the energy intensity of 30 provinces for the five year period between 2004 and 2009. Using the total factor energy efficiency framework, the authors found that in the Eastern region only 33.3% of the total provinces explored had an efficiency level below 0.9, while six provinces in the central region had this level, and so did all the provinces in the Western part. The authors point out that Shanxi is the province with the lowest energy efficiency level in China, and argued further that this result could be linked to the province's abundant energy resources, and lack of any form of energy

efficiency in place. However, the main weakness of the study is the failure to analyse the factors that drive the Chinese industrial sector energy intensity level. This would have made the study more interesting and relevant, by shedding more light to this important part of the debate, especially for a country (China) that is the largest consumer of energy and GHG emitter in the world (EIA, 2015).

In their detailed analysis of the determinants of energy intensity in China, Zeng *et al.* (2014) were able to bring more evidence of the ten year period from 1997 to 2007 to the energy literature. The authors, who employed the use of a structural decomposition analysis (SDA), included the energy intensity, energy mix, sectoral energy efficiency, production structure, final demand structure among sectors, and final demand composition as the decomposition variables. The model used for the analysis is presented below:

$$\Delta e = \tau E_t \hat{r}_t^* L_{d,t}^* y_{sec,t} y_{cat,t} - \tau E_t - \widehat{1r_t^*} - 1L_{d,t}^* - 1y_{sec,t} - 1y_{cat,t} - 1 \qquad (3.27)$$

The evidence from this study suggests that increased efficiency at the sectoral level could explain the reduction in the energy intensity level, while at the same time the rise in exports level brought about by changes in the production structure and demand composition, might account for the increase trend observed between 2002 and 2007 in China. Taken together, these results suggest that the Chinese government may need to structure the exports from the production structure and the demand in a more energy efficient way so to reduce the energy intensity level. However, despite the robust findings of the study, no attempt was made by the authors to use the newly developed cointegration method which would have made the study more reliable.

A panel study of 76 developing countries, by Sadorsky (2013) suggests an that increase in income in these countries reduces energy intensity, while an increase in industrialisation leads to a higher energy intensity level. The author also analysed the impact of urbanisation on energy intensity in the model used for the analysis, and gave three main reasons for the inclusion of the variable. Firstly, is the role played by urbanisation in leading to economies of scale through an increase in the amount of economic activity in urban areas. Secondly, is the increase in demand for energy caused by more motorised traffic as a result of the increase in mobility and transport in the cities. Lastly, is the shift towards more energy intensive goods or sectors, which could be linked to a need for more infrastructure in the urban areas to cater for the higher number of people in cities. However, despite the use of both homogenous and heterogeneous panel models by the author, a mixed result of the impact of urbanisation on energy intensity was found in the study. Furthermore, lack of distinction between the countries to eliminate the lump-together problem, perhaps by income group or other distinct categories, reduced the worth of the study. Also, the use of a panel cointegration technique would have been more suitable considering the unbalanced nature of the panel used for the analysis.

In another study by Adom and Kwakwa (2014), it was found that, using the cointegration technique, the effect of changing technical characteristics of the manufacturing sector in Ghana led to a reduction in energy intensity. Furthermore, the effect of FDI, trade openness and urbanisation were also analysed. They found that energy efficiency was adversely affected by the degree of urbanisation, while a negative impact of trade openness was found, the impact of FDI was positive and insignificant statistically. The importation of more energy efficient machinery through increased economic integration reduced the amount of energy used, while the transfer of new technology through FDI seems not to have improved the energy efficiency level. The impact of FDI needs to be interpreted with caution due to the statistical insignificance of the coefficient in the analysis. However, no explanation was provided by the authors on the role played by income and energy price on the energy intensity level in Ghana.

Furthermore, the impact of energy intensity on carbon emission in 12 SSA was investigated in the study by Shahbaz *et al.* (2015), using the panel cointegration approach. The long run results from the analysis between 1980 and 2012, in the study show that energy intensity has a positive and statistically significant impact on CO2 emissions in the analysed countries.

3.11 Methodological issues raised by the reviewed studies

Many of the empirical studies reviewed in this chapter suffer from shortcomings in the methodological approach adopted by the researchers, type of data used, the reliability of the data sources, and the econometric tests performed. As pointed out by Urban *et al.* (2007), top down models which include the econometric approach used by many of the authors of the studies reviewed, could result in the estimation of incorrect computed outputs elasticities, if an incorrect economic framework is specified and used by the researcher.

The economic framework used for the analysis could be misleading in several ways. For instance, the use of a bivariate model could have led to spurious results in some studies through the bias stemming from the omission of relevant variables (Akinlo, 2008), whereas the specification of a multivariate model for the analysis could have corrected for this. A multivariate model is used for the analysis in this study.

The cointegration technique seems to be a more reliable and rigorous econometric technique for panel data analysis, where multiple countries are analysed. The use of cointegration in empirical analysis requires a longer time period to be efficient. The study uses 34 years in the aggregate demand analysis, where the cointegration method is employed, whereas, some of the reviewed studies used a shorter time period (see for example, Akinboade *et al*., 2005; Iwayemi *et al*., 2010; Mensah, 2014 in Table 3.1 above).

Some of the reviewed studies also neglected the time series properties of the variables and their order of integration. In other words, where a mixed order of integration was found among the variables, the methodological approach employed by some of the studies did not take this into consideration. This important point is taken into account in the econometric approach chosen and used in this study. The Pooled Mean Group (PMG) technique is used in the analysis, where a mixed order of integration is found in this research. PMG is a panel extension of the Autoregressive Distributed Lag (ARDL) model popularly used in time series analysis due to its ability to estimate a cointegration relationship among variables even when they have a different order of integration (Martins, 2006; Morshed, 2010; Frimpong and Adu, 2014).

Likewise, most of the reviewed studies in SSA seem to have neglected some important variables like the degree of urbanisation and economic structure, both of which are important when considering the current trends and dynamics in SSA (Mensah, 2016). These two variables will be included in the cross-country econometric model analysed in this study.

In view of these limitations, the study attempt to overcome the shortcomings identified in the reviewed literature by taking into account the characteristics of a developing country for the variable specification of the model used, and also by using a comprehensive data set from reliable data sources, employing the use of the cointegration technique, linear panel models and performing all the necessary econometric tests.

3.12 Chapter summary

The purpose of this chapter was to discuss the theoretical framework for the study, and also to identify the relevant variables to be used for the analysis of energy demand in Sub-Saharan Africa (SSA). This was achieved through the critical review of the previous studies on the subject, especially those in developing countries.

Most of the studies at the aggregate level concluded that income and price variables are the main determinants of energy demand. This is in line with economic theory of the law of demand. Changes in the economic structure, the degree of urbanisation and the price of substitutes are also important factors that influence total energy demand, as evidenced in the results of some of the studies reviewed in this chapter.

The choice of the variables to be used for the disaggregated energy demand study depends on the sector under study, and other variables are added to the price and income variables in the analysis, while at the same time the particular variables chosen depend on their relevance to the sector to be analysed as well as data availability.

Since the seminal paper by Kraft and Kraft (1978) and other subsequent studies, there are conflicting and mixed results on the true direction of the causality between economic growth and energy consumption. These include studies on carbon emissions as a variable in the model. However, the review of the research on the determinants of energy intensity shows that it may have a role to play in the energy security issues, and to mitigate climate change in developing countries.
The next chapter discusses the methodology to be used for the analysis in this study.

Chapter 4 : Econometric Methodology

4.1 Chapter overview

This chapter provides an account of the methods used for the estimation of panel data models, which is the main methodological framework used for the empirical analysis in this PhD study. The definitions, estimations and interpretations of panel data models are discussed as adopted for the estimation of the aggregate and disaggregated energy demand analysis.

The chapter is structured as follows. First, the definition of panel data is introduced in the first section, in addition to a discussion of limitations and benefits of using panel data. In the sections that follow, the panel methods discussed are grouped under two main headings: panel linear models, and non-stationary linear models.

Under the panel linear models, the fixed effects and random effects models that take into account the unobserved differences in the countries analysed in Sub-Saharan Africa (SSA), are explored. The Prais-Winsten (PW) regression model is explained in the last sub-section. Under the other division of panel models, the issues of stationarity and unit roots are explained, a basis which is then used for the discussion of panel unit root tests, before moving on to the discussion of panel cointegration. The chapter concludes with the motivation and rationale for the chosen methodology for the empirical analysis.

4.2 Introduction to panel data

Panel data are a type of data that includes, in addition to the temporal element, the cross section of observations of economic variables over time (Baltagi, 2008, p. 1). The definition suggests that panel data entail cross section as well as time series components. The Sub-Saharan African energy demand panel dataset, the dataset used in this PhD study, gives us an ample opportunity for analysing the driving forces behind energy demand between 1980 and 2013 in Sub-Saharan Africa by building an econometric model for the analysis. Panel data may be classified depending on the number of the ‘group’, into micro panels and macro panels. Group is defined in the econometric literature as the observation units in a panel dataset. A micro panel data set is collected for a large group over a short time period. While, the second classification of panel data comprises of a small number of groups and it is observed annually over a relatively long time period, between 20 and 60 years (Baltagi, 2008).

The Sub-Saharan African (SSA) energy demand model is a macro panel dataset, because it includes a cross section of different economic variables of the studied countries in the region over 33 years. The difference between the two types of panel data lies on the relative number of N and T. Macro panel data sets have a moderate N size (for example, aggregated into a region or sub-regions) and a large T (long time period) dimension. In general, macro panel datasets tend to suffer from the nonstationarity issue in the time series (at least in some variables), like unit roots, as well as structural breaks, a problem that arises because of the long time series dimension.

Cross-country dependence is another issue to be considered when estimating a macro panel dataset, or a panel data that has a large cross section and time series dimension. Failure to account for this might lead to making wrong or misleading inferences (Baltagi, 2008). Chudik and Pesaran (2013) provide a survey of the techniques that have been proposed in the literature for overcoming the problem of cross dependence of errors. In their study, Chudik and Pesaran (2013) evaluated techniques such as, the factor error structure employed in cross-sectional dependence, principal components techniques, the common correlated effects approach, and the quasi maximum likelihood estimator, among others.

Despite the few limitations of panel data mentioned in the paragragh above, the use of panel data offers many advantages over studies that use conventional data sets such as time series or cross sectional data. The use of nonstationary panel data enables the combination and derivation of the advantages of both cross section and time series data. That is, to get the best of both worlds in terms of the availability of more data and power across the cross section and also the advantages that comes with removing unit roots in time series data.

According to Hsiao (2014), there are several benefits of using the panel data approach over the traditional time series or cross sectional data tehniques. Some of the advantages highlighted by Hsiao (2014) include: panel data have a large dataset, which makes it more informative with high degrees of freedom and reduced collinearity among the regressors, thereby improving the reliability of the econometric estimations of the model; panel data are more suitable for studying the dynamics of adjustment than cross sectional data; panel data reduce the impact of omitted or unobserved variables in the results of the model; panel data enable the researcher to model and test more complicated behaviour models, thus allowing for tests of various levels of heterogeneity in the sample and; the unit root tests employed in macro panels have standard asymptotic distributions unlike the case of traditional time series data, which is known to have nonstandard distributions.

It seems evident from the discussion above that special techniques have to be employed when using a panel dataset in an empirical study, so as to estimate the parameters in the model correctly. The section that follows discusses the panel data regression models and how they are estimated and specified for econometric analysis.

4.3 Panel linear models

In this section, the specifications, definitions and interpretations of linear panel models are discussed. The fixed effects model, random effects model and PW models work best in a balanced panel. Since the disaggregated energy demand dataset used in this study is a balanced panel, these models will be employed for the estimations. Therefore, in the sub-sections that follow, each of these models is discussed one after the other. Relevant tests used to validate the reliability of estimated coefficients are also discussed.

4.3.1 Panel data regression model

A regression model analyses the relationship between a dependent variable and one or more explanatory variables. For instance, if z is the dependent variable and y is a vector containing different independent variables, the regression equation can be written as:

$$z_{it} = \alpha + \beta y_{it} + u_{it} \quad (4.1)$$

where i represents i, 2,3,…..,N and t stands for 1,2,3,…T

i is the cross-sectional units while T is the total time period that corresponds to each cross sectional unit and u_{it} is the 'white noise' error term, expected to be normally and independently distributed with zero mean and constant variance.

We can derive the expected value of z, and represent it as:

E (z/y) = $\alpha + \beta y$

The OLS estimator can be used to estimate the regression model above. However, in order not to have an unbiased estimate some conditions must be met. Some of the assumptions which must be met were highlighted by Gujarati and Porter (2009) and Al-Kuwari (2007), and discussed briefly below:

a) Error term: The disturbance term mean or expected value is zero given the value of y, that is

$$E(u_{it}/y) = 0$$

b) Homoscedasticity: This condition states that the variance of the error term is constant regardless of the value taken by y, that is,

$$\text{var}\,(u_{it}/y) = \sigma^2$$

c) Linearity: The regression model in equation (1) above is linear in parameters, although the variables themselves may be nonlinear in nature.
d) Normality: The error term should be normally distributed. That is $u_{it} \sim N(0, \sigma^2)$.
e) Serial correlation: There should be no autocorrelation in the error term in the model. That is,

$$\text{cov}(u_{it,} u_{i(t-1)}) = 0$$

f) Degrees of freedom: In a regression model, the number of observations must be greater than the estimated parameters.

g) Multicollinearity: There should be no perfect linear relationship among the independent variables in a model. An assumption usually verified by computing a correlation matrix for all the variables in the model; as a rule of thumb, a correlation higher than 0.75 would display a concerning level of collinearity, while correlation values below 0.75, would indicate relatively low and hence innocuous levels of collinearity.

h) Zero covariance: The error term and the explanatory variable should have a zero covariance, that is,

$$\text{cov}(u_{it}, y_{it}) = 0$$

If the stated assumptions hold, OLS is said to be the best linear unbiased estimator (BLUE) otherwise the OLS estimate is biased if any of them is violated. Therefore, it is important to perform all necessary statistical tests to ensure that all conditions of OLS are met in empirical analysis. However, even if all the conditions of OLS are met, OLS may not be BLUE if there is a considerable variability in the dependent variable analysed (Gujarati and Porter, 2009, p. 371). In a dataset with wide variability in the dependent variable, GLS (Generalised Least Squares) may be more suitable for the analysis. In this section, we have discussed the pooled OLS used in panel data and its properties. In the sections that follow, the static panel models are discussed.

4.3.2 Static linear panel model

The static linear panel models include estimators such as the fixed effects model and the random effects model, among others. Each of these models has a different specification, assumptions and suitability. In order to explain the differences between these models, the time series linear model and the cross sectional linear model are used to build a panel model employed for the illustration.

First, let us examine a cross-sectional model of the form

$$z_i = \alpha + \beta y_i + u_i \tag{4.2}$$

where z_i represents the value of a cross-sectional unit, for example a country at a single point in time. y_i is a vector of explanatory variables, while u_i denotes the cross-sectional errors.

Next, let us examine a time series model

$$z_t = \alpha + \beta y_t + u_t \tag{4.3}$$

In equation (4.3), z_t stands for the value of z at time period t, y_t is a vector of independent variables and u_t is the error assumed to be white noise.

Both models presented in equations (4.2) and (4.3) above are based on the assumption that the fixed parameters $\alpha \; and \; \beta$ accounts for the effect of z on y. However, in panel data which is a combination of both models is based on the assumption that the factors that account for the impact of z on y in each cross-sectional unit of the model could be different (Al-Kuwari, 2007).

Therefore, following the assumption that $\alpha \; and \; \beta$ may vary over the cross-sectional units over time, equations (4.2) and (4.3) become

$$z_{it} = \alpha_{it} + \beta_{it} y_{it} + u_{it} \tag{4.4}$$

with i ranging from 1 to N, and t from 1 to T.

It should appear evident from equation (4.4) above that it might be difficult to estimate the model because the number of parameters (NT (K+1)) is greater than the degrees of freedom of the model (NT). This condition requires that we impose a structure on (4.4) so as to be able to estimate the model. Al-Kuwari (2007) pointed out that one of the ways to achieve a valid model for equation (4.4) is to assume that β (known as the structural parameter) is identical across all cross-sectional unit over time, and introduce a new variable known as the incidental parameter.

In other words, the incidental parameter is the additional structure imposed on model (4.4) to be able to estimate the model. This suggests that the incidental parameter accounts for the differences (heterogeneity represented here by δ) across countries over the period of study, which could be omitted as an explanatory variable. These incidental parameter may include omitted variables that are stable over time, individual in-variant and varying variables. On one hand, individual invariant variables are variables that are similar for all i at a particular point in time but change through time. This includes variables such as prices and the rate of interest. On the other hand, time varying variables are variables that are different across both cross-section and time at a given period. Examples of such variables include sales volume and profits. In the empirical literature, the impact of the heterogeneity across the cross-sectional units is said to be either, random, fixed or mixed.

Therefore, the main task of the researcher in empirical analysis is to obtain an unbiased estimate of the structural parameter by dealing with the unobserved differences correctly. Having discussed how the heterogeneity came about and how it can be modelled to be (either, random, fixed or mixed effects), in the sections that follow, the random effects model and the fixed model are discussed in turn. How the appropriate model is chosen is also discussed in the context of the Hausman test.

4.3.3 The random effects (RE) model

In regression analysis, the standard assumption is that all the factors that influence the dependent variable but are not included as explanatory variables, are captured by the error term. In the RE model the unobserved heterogeneity is assumed to be based on random factors that are independently and identically shared among the individual units of the panel (Verbeek, 2012). Therefore, we write the RE model as:

$$z_{it} = \alpha_{it} + \beta y_{it} + \mu_i + u_{it} \tag{4.5}$$

where $u_{it} \sim IID(0, \sigma_u^2)$; $\mu_i \sim IID(0, \sigma_\mu^2)$

In equation (4.5), $\mu_i + u_{it}$ is categorized as an error term which has two main divisions. The first division does not vary over time and it is known as the individual specific component, while the other part of the error term is assumed to have no correlation over time. The basic assumption in the model is that the unobserved differences across i are not correlated with y, and they are distributed randomly with a common mean (Baltagi, 2008, p14). In other words, $both\ \mu_i\ and\ u_{it}$ are mutually independent and they have no relationship with either the time series or the cross-sectional components. That is explained using the equations below:

E $(\mu_i\ u_{it}) = 0$ for all i and t

E $(\mu_i\ \mu_K) = 0$ $\quad$ (i$\neq k$)

E $(u_{it}\ u_{ij}) = E(u_{it}\ u_{kt}) = E\big(u_{it}\ u_{kj}\big) = 0$ $\qquad$ (i$\neq k$; k $\neq j$)

The Breush and Pagan (1979) Test

In order to test whether the presence of random effects is present in a panel, a Lagrange Multiplier (LM) test was developed by Breusch and Pagan (1979). The LM test statistic is

$$\text{LM} = \frac{NT}{2(T-1)}\left[\frac{\Sigma(\Sigma\, u_{it})^2}{\Sigma\Sigma\, u_{it}^2} - 1\right]^2 \sim x^2(1)$$

u_{it} represents the pooled OLS error term.

The null hypothesis of the test is that cross-sectional variance units are zero. Likewise, under H0, x^2 has one degree of freedom and it is asymptotically distributed. Therefore, the rejection of the null hypothesis means that random effects model is acceptable.

The use of the random effects model is useful when the sample used is drawn randomly from a population and the chosen sample is representative of the population under study (Baltagi, 2008, p.14). The advantages of using a random effects model are: a) when the sample size increases, the number of parameters to be estimated does not change; b) The derived estimator (known as the generalised least squares) is efficient because it applies both the within and between variations; and c) the effect of a variable that does not change over time across cross-sectional units can be estimated. However, despite the usefulness of the model its strong assumption of independence of the regressors from the unobserved heterogeneity seems to be too strong (Al-Kuwari, 2007).

4.3.4 The fixed effects (FE) model

In a fixed effect model, the intercept in the regression equation varies across the cross-sectional units. The fixed effect model is represented by:

$$z_{it} = \alpha + \beta y_{it} + \mu_i + u_{it} \tag{4.6}$$

In the model above, the unobserved heterogeneity δ has two main parts: α and μ_i. The common fixed effects across the cross-sectional units i represented by α is the first part. While the deviation from the common effect which is constant for each individual is represented by $\mu_{i,}$ is the second part of the unobserved heterogeneity. The underlying assumptions of the fixed effects model are:

$$u_{it} \sim N(0, \sigma_u^2)$$
$$\mathrm{E}\,(u_{it}\, u_{ij}) = E(u_{it}\, u_{kt}) = E\big(u_{it}\, u_{kj}\big) = 0 \qquad (\mathrm{i} \neq k; \mathrm{k} \neq j)$$

That is, y_{it} are uncorrelated with all u_{it}. As stated earlier that the intercepts varies across individuals, we can use dummy variables to account for this in each individual units i represented by:

$$z_{it} = \sum_{j=1}^{N} \alpha_j d_{ij} + \beta y_{it} + u_{it} \tag{4.7}$$

In (4.7) above, we have N dummy variables in the model if d_{ij}=1 and i= j and 0 elsewhere. Then, OLS could be used to estimate the model parameters $\alpha_{1,} \alpha_{2,} \alpha_{3,\dots\dots\dots,} \alpha_{N\, and}\ \beta$. However, it may not be appealing to estimate a model with many regressors. If so, the regression can be estimated using deviations from individual means. The transformation from individual mean is presented by:

$$z_{it} - \bar{z}_i = (y_{it} - \bar{y}_{it})^I \beta + (u_{it} - \bar{u}_i) \quad (4.8)$$

The results of both models in (4.7) and (4.8) above will be identical (Verbeek, 2012). The results of the model in (4.8) for β is called the fixed effects or 'within estimator'.

In essence, the fixed effects model accounts for the differences within individual units in a model. However, the model allows some degree of correlation between the explanatory variables and the units analysed. Despite its usefulness, there are some drawbacks with the use of the model. The main shortcoming associated with the model is that the number of unknown parameters increases if the number of observations increases.

4.3.5 Random effects versus fixed effects models

From the discussion above, it is evident that both fixed and random effects models can be used for panel data estimation. But then, we need to choose the appropriate one to use given the nature of our sample. The main idea is whether to assume that individual specific effect (μ_i) parameter is a variable randomly drawn from the population or to estimate it as a fixed variable. The choice of which of the two models to use has generated a considerable debate in both the statistical and biometrics literature, which has now extended into the panel data econometrics literature (Baltagi, 2008, p.19). Usually the RE model is preferred when dealing with census data rather than a sample drawn from a wider population. However, basic intuition might suggest that if the datasets represents countries or states, then μ_i might be treated as a fixed parameter because the variable is not from a random selection from the distribution. Likewise, if the panel datasets represent variables such as individuals or firms then μ_i is likely to be estimated as a random component of the model. Based on the illustration, the main point is to ascertain if there is correlation between the observed variables (z) and individual specific effect μ_i.

Therefore, the fixed effects model is employed if there is correlation between y and μ_i and it also follows that the random effects model is used if there is no correlation among them. This suggests that fixed effects model is only BLUE (best linear unbiased estimator) if y and μ_i are correlated. To know if there is any correlation between y and μ_i, Hausman (1978) proposed a test. The test statistic is given by:

$$H= \left(\hat{\beta}_{FE} - \hat{\beta}_{RE}\right)^{l}\left[cov\left(\hat{\beta}_{FE}\right) - cov\left(\hat{\beta}_{RE}\right)\right]^{-1}\left(\hat{\beta}_{FE} - \hat{\beta}_{RE}\right) \tag{4.9}$$

In (4.9) above, $\hat{\beta}_{FE}\ and\ \hat{\beta}_{RE}$ are vectors of coefficients estimates from the fixed effects and the random effects model, respectively. It excludes the coefficients from time dummies and time-invariant variables. The statistic (H) is asymptotically distributed as x^2 with degrees of freedom equal to the sum of the coefficients in $\hat{\beta}_{FE}\ and\ \hat{\beta}_{RE}$.

The null hypothesis of the test is:
H0: The fixed effects estimators and the random effects estimators do not differ substantially. In other words, z and μ_i are not correlated.
The alternative hypothesis of the test is:
H1: The fixed effects estimator and the random effects estimators differ substantially.

After either of the models is selected for the estimation, it is important to ensure that the estimate is BLUE (Best Linear Unbiased Estimator). In the event that this condition is not met, as it was the case in some of the models analysed in this study, where evidence of heteroscedasticity was found (see Chapter 6), the Prais-Winsten regression with panel-corrected standard errors can be employed (see for example Poumanyvong and Kaneko, 2010; Knight and Schor, 2014) to obtain unbiased estimates.

Having discussed static panel models used in balanced panels in the sections above, the methods used to estimate non-stationary panels are discussed in the sections below. Issues that could be raised by the use of macro panel data, solutions to the problems of nonstationarity and cross dependence are also explored.

4.4 Nonstationary panel models

In this section, the methods used to estimate non-stationary panel are discussed. However, because of the nature of the time series component element in a macro panel, it is critical to properly address the issues of unit roots, structural breaks, cointegration and cross dependence.

4.4.1 Econometric concept of stationarity and unit roots

Before detailing the panel unit root tests commonly used in the literature, this section begins by offering a clear explanation of two econometric concepts which are essential to provide adequate background to the discussion that will follow. That is the issue of stationarity and of unit roots.

The issue of stationarity in times series data, is analogous to that in panel data. When a variable is non-stationary, it is said to have unit root problem (see Figure 4.1). That is, the variable comes from a stochastic distribution, with either or both of its mean and variance changing over time. A variable could be said to be stationary, if both its mean and variance do not change over time (see Figure 4.2), and the covariance does not depend on the time, rather on the distance between the two time periods it was computed (Gujarati and Porter, 2009).

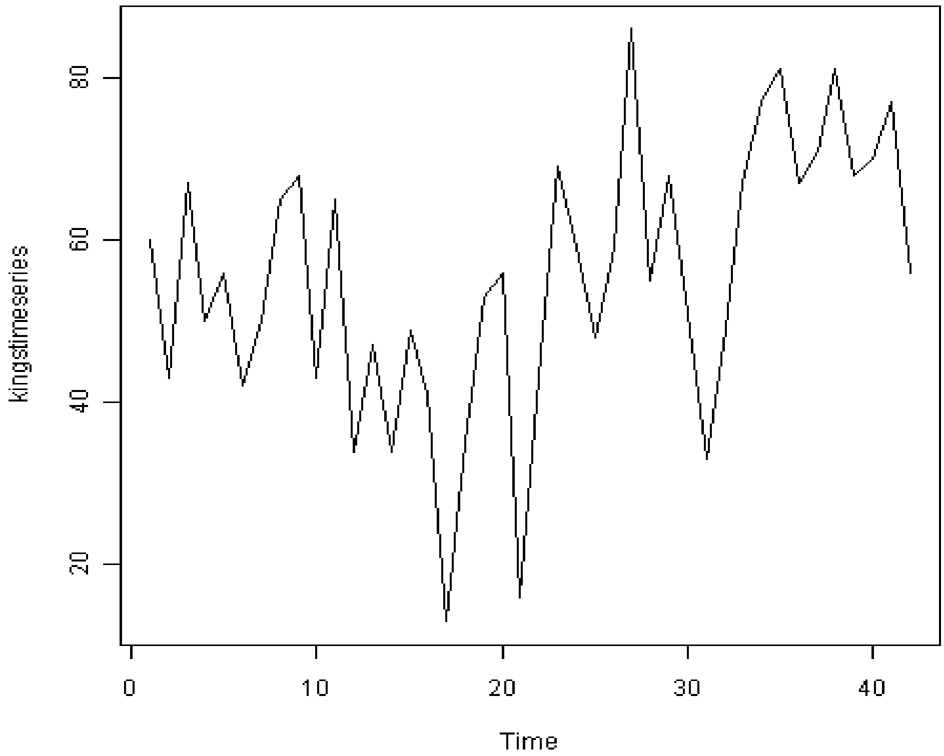

Figure 4.1: A non-stationary time series

(**Source:** Coghlan, 2010)

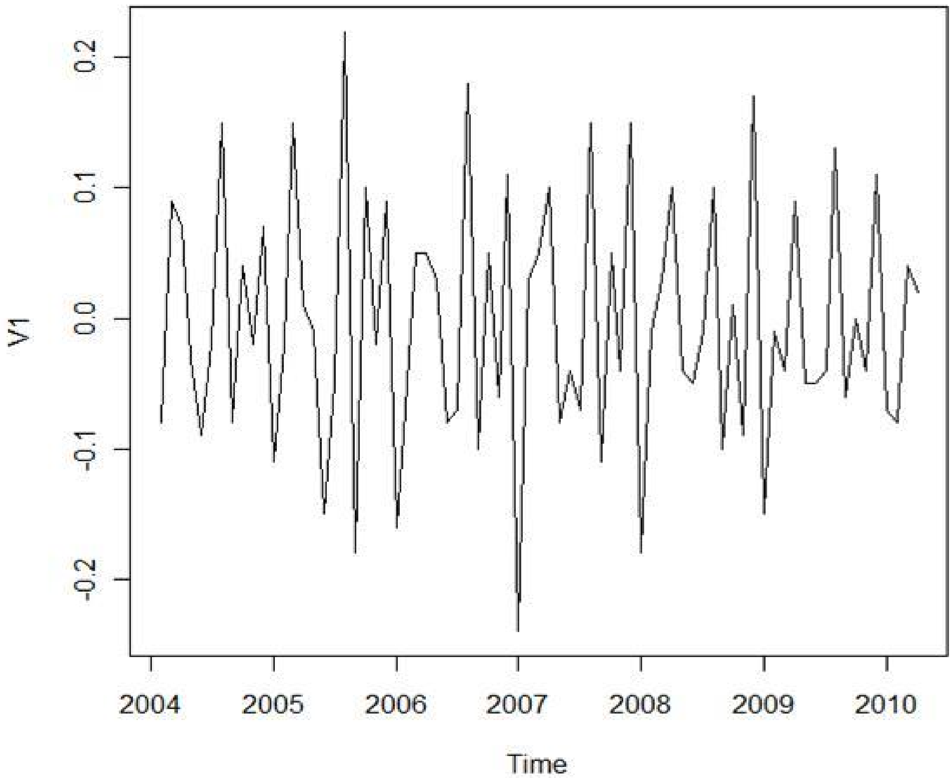

Figure 4.2: A stationary time series

(**Source:** Coghlan, 2010)

Gujarati and Porter (2009, p.740) used the three equations below to explain the concept of stationarity in a time series model.

$\mathrm{E}(Y_t) = \mu$ (4.10)

$\mathrm{var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2$ (4.11)

$\gamma_x = E[(Y_t - \mu)(Y_{t+x} - \mu)]$ (4.12)

where equations (4.10) and (4.11) represent the mean and variance of Y, respectively, and (4.12) is the covariance at lag *x*, that is between two values of Y (Y_t and Y_{t+x}), the covariance is γ_x. Also, if the three equations are not true for a variable, then the series is non-stationary. However, there is a weak stationarity if the variance changes but the mean remains the same over time.

To illustrate non-stationarity or the presence of a unit root in a series, let us consider the AR(1) model below:

$y_t = \alpha y_{t-1} + u_t$ (4.13)

In other words, if the changes in the value of y are drawn from a stochastic distribution, which would suggest that it is from u_t, then the process is non-stationary, and definitely contains a unit root (Asteriou and Hall, 2011, p.231).

Two different values which could be taken by α are used to explain the stationarity concept using backward substitution. That is:

If α=1 in (4.13) above, then

$y_t = y_{t-1} + u_t$ (4.14)

Using backward substitution, we have

$y_1 = y_0 + u_1$ (4.15)

$y_2 = y_1 + u_2 = y_0 + u_1 + u_2$ (4.16)

$\vdots \quad \vdots \quad \vdots$

$y_t = y_0 + \sum_{i=1}^{t} u_i$ (4.17)

y_t has an infinite memory in a random walk model, which could be seen from figure 1 above.

Now, let us consider a case where $\alpha< 1$ in (4.13) above, then

$$y_t = \alpha y_{t-1} + u_t \tag{4.18}$$

Using backward substitution, we have:

$$y_1 = \alpha y_0 + u_1 \tag{4.19}$$

$$y_2 = \alpha y_1 + u_2 = \alpha^2 y_0 + \alpha u_1 + u_2 \tag{4.20}$$

$$y_3 = \alpha y_2 + u_3 = \alpha^3 y_0 + \alpha^2 u_1 + \alpha u_2 + u_3 \tag{4.21}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$y_t = \alpha^t y_0 + \sum_{i=0}^{t-1} \alpha^i u_{t-i} \tag{4.22}$$

y_t has a finite memory because as T tends towards infinity, $\alpha<1$ and $y_0=0$, also $\sum_{i=0}^{t-1} \alpha^i u_{t-i}$ tends toward zero (0).

To enable the researcher to choose the appropriate technique for an economic model, the issue of nonstationarity must be considered. And the use of conventional regression analysis which was founded on the basis that both the means and covariances of most economic variables are independent of time, as seen in the case above when $\alpha<1$, could, therefore, be misleading. In other words, most of regression analysis is modelled under the assumption that the variables come from a stationary process which, in fact, is rarely the case in practice (as a result of stochastic trends in economic data due to technology, structural breaks, etc.).

Failure to account for the nonstationarity of economic variables in regression analysis gives rise to the spurious regression problem. The issue of spurious regressions was first noticed by Yule (1926), who observed that in models where Ordinary Least Squares (OLS) regressions are employed, the results could suggest a statistically significant relationship even when none exists among the variables. In his study, the author investigated the problem of 'nonsense' correlations between individual time series when they are in levels and integrated of order one.

Yule (1926) was the first to suggest that if the non-stationary properties of the model are not taken into account, the model could give biased results and hence wrong correlations. Bispham (2005) pointed out three main properties of spurious regressions: (a) inconsistency in the estimates; (b) divergence of both the OLS t and F-statistics; and (c) residuals may not tend towards zero.

Further evidence on the spurious regression problem was provided in the study by Granger and Newbold (1974) who popularised the problem by showing that the problem could be identified (diagnosed) from results which simultaneously display an overly high R-square measure (the coefficient of determination which measures 'goodness of fit'), statistically significant t- statistics, but with residuals recording a high serial correlation. These features could also explain why the logarithm of data is typically taken in econometric analyses, as most of the series are trended. For example,

$$y_t = 1.1y_{t-1} \qquad (4.23)$$

Taking the log of (4.23), we have:

$$\text{Log}(y_t)=\log(1.1)+\log(y_{t-1}) \qquad (4.24)$$

To make the variable stationary, we need to take the first difference. Then (4.23) becomes integrated of order 1, that is I(1). On the other hand, a variable will be integrated of order 2, if it was differenced twice to become stationary.

The problem of nonstationarity in the United States, was first modelled by Nelson and Plosser (1982) using four US macroeconomic time series. Their research led to other methodological and empirical studies, where the unit roots in time series data have been tested. In the section below (4.4.2), the unit root test which was implemented by Nelson and Plosser (1982) and by several other researchers in the econometric literature like Dickey and Fuller (1979, 1981), is discussed.

4.4.2 The Dickey-Fuller unit roots test for time series data

The first unit root test strategy for time series data was developed by Fuller (1976) and Dickey and Fuller (1979, 1981). In order to develop the model, constants and trends are added to the regression in a step by step sequence and the existence of unit roots is used to ascertain non-stationarity.

Consider the AR(1) model

$$y_t = \alpha y_{t-1} + u_t \tag{4.25}$$

The null hypothesis is $H_o : \alpha = 1$, against the alternative $H_1 : \alpha < 1$. Where $u_t \sim IID(0, \sigma^2)$.

Under the null hypothesis that the variable has a unit root, the data generating process (DGP) will be:

$$y_t = y_{t-1} + u_t \tag{4.26}$$

However, under the alternative hypothesis of the series being stationary, α is from a normal distribution and is given by

$$\sqrt{T}\ (\hat{\alpha} - \alpha) \xrightarrow{d} N(0, 1 - \alpha^2) \tag{4.27}$$

Where $\xrightarrow{d}$ in the equation, means that the series converges in distribution. Also, in the case of the null hypothesis, the estimator denoted by $\hat{\alpha}$ which stands for super consistent, means that the distribution coverges (the rate of T as against$\sqrt{T}$) in stationarity.

The functional central limit theorem (FCLT) can be used to know the limiting distribution under the null hypothesis, as T→ ∞, in equation (4.25):

$$\mathrm{T}\ \widehat{(\alpha - 1)} \Rightarrow \frac{\left(\frac{1}{2}\right)[W(1)]^2 - 1}{\int_0^1 [W(r)]^2 dr} \tag{4.28}$$

W(1) and W(r) stand for standard Brownian motions[1], while the arrow sign ($\Longrightarrow$) means that the system converges weakly. Monte Carlo simulations are used to generate the critical values, which are available from the set of tables provided in Fuller (1976) p.371 case $\hat{\rho}$.

The t-statistic under the null hypothesis, $H_o : \alpha = 1$, and the limiting distribution for the AR (1) process as T→ ∞, is given by:

$$t_\alpha \Longrightarrow \frac{\left(\frac{1}{2}\right)\{[W(1)]^2 - 1\}^{1/2}}{\int_0^1 [W(r)]^2 dr} \qquad (4.29)$$

Likewise, to obtain the critical values for the distribution, Monte Carlo simulations are used (see Fuller, 1976, p.373), for the case $\hat{\tau}$.
(4.25) can be rewritten as:

$$\triangle\, y_t = \rho y_{t-1} + u_t \qquad (4.30)$$

ρ= 1- α. Hence, under the null hypothesis, since $\alpha = 1$, ρ=0. Therefore, a t-test of ρ= 0 can be used to test for a unit root in y_t. Asteriou and Hall (2011, pp. 342-343) provide - by way of illustration - a summary of how the simple Dickey-Fuller test for unit roots is performed.

The unit root test discussed so far, assumes that the error term is white noise (that is u_t is $IID(0, \sigma^2)$), which might not always be the case. So, to cater for autocorrelation in the error term, one can add an extra lagged variable of the dependent variable. Dickey and Fuller used two approaches to solve the autocorrelation issue in the error term. Both methods used are explained in turn.

The first approach involves the addition of the lagged differences of the dependent variable in the form of parametric corrections in the regression model to eliminate the serial correlation of u_t. Fuller (1976, p.374) used the equations below to explain the process as cited in Bispham (2005, p.8).

[1] Brownian motion or a Wiener process is defined as a real valued stochastic process

$$y_t = \mu + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_\psi y_{t-\psi} + u_t \quad (4.31)$$

The 'augumented' Dickey-Fuller (ADF) test equivalent of (4.31) will be:

$$\Delta y_t = \mu + \psi y_{t-1} + \sum_{j=1}^{\psi-1} \alpha_j \Delta Y_{t-j} + u_t \quad (4.32)$$

Where, $\psi = (\beta_1 + \beta_2 + \beta_3 + \cdots \ldots\ldots + \beta_{\psi-1} - 1)$.

Therefore, the ADF regressions are employed when the error terms are serially correlated.

The second approach was developed by both Phillips and his colleague in Phillips and Perron (1988), in which a nonparametric technique was used to correct for the autocorrelation of the error term. In this test, the t-statistic of the ψ cofficient is corrected so as to 'whiten' the residuals in the series. The asymptotic distributions of the two approaches are the same (Asteriou and Hall, 2011, p.297).

Using the case of a model with a unit root, the Phillips and Perron (1988) test shows that:

$$Z_\alpha = T(\hat{\psi} - 1) - \frac{1}{2}(T^2 \hat{\sigma}_\psi^2 \div s^2)(\hat{\lambda}^2 - \hat{\Upsilon}_0) \Rightarrow \frac{\left(\frac{1}{2}\right)\{(W(1))^2 - 1\}}{\int_0^1 (W(r))^2 d_r} \quad (4.33)$$

$\hat{\Upsilon}_j = \frac{1}{T}\sum_{t=j+1}^{T} \hat{u}_t \, \hat{u}_{t-j,} \hat{u}_t$ represents the OLS sample residual derived from the regression that is being estimated, $\hat{\lambda}^2 = \hat{\Upsilon}_0 + 2\sum_{j=1}^{a}(1 - \frac{j}{(q+1)}) \hat{\Upsilon}_j$, $s^2 = \frac{1}{(T-K)}\sum_{t=1}^{T} \hat{u}_t^2$, k stands for the number of parameters estimated in the regression, $\hat{\sigma}_\psi$ is ψ standard error for the OLS. The critical values used in the Dickey-Fuller test without serial correlation is applicable for the test.

Next, for the ADF case in a model with a unit root, the distribution is given by:

$$\frac{T(\hat{\psi} - 1)}{1 - \hat{\alpha}_1 - \hat{\alpha}_2 - \cdots .. - \hat{\alpha}_{\psi-1}} \Longrightarrow \frac{\left(\frac{1}{2}\right)[[W(1)]^2 - 1]}{\int_0^1 [W(r)]^2 dr} \quad (4.34)$$

$\alpha_{1,}\ \alpha_{2,}\ \alpha_{3,}\ldots\ldots\ldots\alpha_{,\psi-1}$ are identical to those of the (4.32) above, and the same critical values are employed.

In the sub-section above (4.4.2), the Dickey-Fuller test for unit root testing in time series data was explored, and it should be noted that the choice of technique chosen by the researcher will depend on the Data Generation Process (DGP) of the series in a model. Bispham (2005) concludes that depending on the number of differences taken in a series, when testing for unit roots, the steps to be taken when testing for unit roots are as follows: Plot the data in both levels and differences in a graph, and then the autocorrelations of the data in differences should also be graphed. These steps should be taken for all differences of the data taken. In this research study, the graphs of the series used in the analysis will be presented.

4.4.3 Panel data units root tests

In the literature, various tests have been proposed for testing the presence of unit roots in panel data series. The testing procedure is more complex than that employed in testing for nonstationarity in the context of time series data only. However, most of the tests used are an extension of the ADF tests (Asteriou and Hall, 2011, p.297). Some of the tests used include: Levin and Lin's test, Lin and Chu's test, Pasaran and Shin's test, Maddala and Wu's test, Hadri's test and Breitung's test. Some of these tests are discussed below.

4.4.4 The Levin, Lin and Chu (2002) panel unit root test

Levin and Lin (1992) developed one of the first unit root tests for panel data, which was presented as a working paper in 1992. The full paper was later published by Levin *et al.* (2002) who proposed the use of a panel unit root test against doing the test for each of the cross-sectional units. Based on three main assumptions for a panel comprising a cross section of individuals (i) and time series observations (t) in the stochastic process (y_{it}):

Assumption (a): The stochastic process (y_{it}) is generated from one of the three models

Model 1 $\Delta y_{it=} \psi y_{it-1} + \varrho_{it}$

Model 2 $\Delta y_{it=} \alpha_{0i} + \psi y_{it-1} + \varrho_{it}$

Model 3 $\Delta y_{it=} \alpha_{0i} + \alpha_{1it} + \psi y_{it-1} + \varrho_{it}$

In all the models, -2< $\psi \leq 0$, in all i= 1,…,N.

Assumption (b): The white noise process is independently distributed across individuals and it follows as an invertible stationary ARMA process for all cross-sections.

$$\varrho_{it} = \sum_{k=1}^{\infty} \theta_{ik} \varrho_{it-k} + u_{it}$$

Assumption (c): For all cases where i=1, 2,3,…,N and where t=1,2,3,…T

E($\varrho_{it}^4) < \infty; E(u_{it}^2) \geq B_u > 0; and\ E(\varrho_{it}^2) + 2\sum_{k=1}^{\infty} E(\varrho_{it}\varrho_{it-k}) < B_\varrho < \infty$

In model 1, y_{it} possess an individual-specific mean and we have the null hypothesis H_0: ψ=0, against the alternative hypothesis H_1: ψ<0. For model 2, because the series has no time trend, we have H_0: ψ=0 and α_{0i}=0 across all individual i, against H_1: ψ<0 and $\alpha_{0i} \in R$. However, in model 3 with both an individual specific mean and time trend, H_0: ψ=0 and α_{1i}=0 and the alternative hypothesis H_1: ψ<0 and $\alpha_{1i} \in R$.

Using the general form hypothesis in (4.35), the authors proposed a three step procedure for testing for unit roots in panels if the lag order in a series is unknown. Firstly, for each of the cross sections a different augmented Dickey-Fuller (ADF) regression is carried out. Secondly, the ratios of the short and long run standard deviations are computed. Finally, the panel test statistics is computed.

$$\Delta y_{it} = \zeta y_{it-1} + \sum_{L=1}^{P_i} \theta_{iL} \Delta y_{it-L} + \alpha_{mi} d_{mt} + u_{it}, \qquad \text{m=1,2,3} \qquad (4.35)$$

d_{mt} represents the deterministic variables vector, α_{mi} is the model's (m=1.2 ,3) vector coefficients, the null hypothesis is that the time series is non-stationary against the alternative hypothesis of stationarity. A summary of the steps for performing the Levin, Lin and Chu test is given in Baltagi (2008, p 276). Levin *et al.* (2002) recommend the test for a panel with time series (T) between 25 and 250, and a cross-sectional size (N) between 10 and 250.

Despite the merits of the test proposed by Levin *et al.* (2002), it would have been more robust if it could be used for a panel where there is cross-sectional correlation. Also, the assumption of unit root across all cross-sections reduces its applicability.

4.4.5 The Im, Pesaran and Shin (2003) panel unit root test

Im *et al.* (2003) (henceforth, IPS) addressed the drawbacks of the test proposed by Levin *et al.* (2002) by allowing for heterogeneity across the individuals in the series.

Using a first order autoregressive (AR1) process

$$\Delta y_{it} = \alpha_i + \delta_i y_{it} + u_{it} \qquad (4.36)$$

where i=1,2,3,….,N and t=1,2,3,….,T, u_{it} represents the error term and it has different serial correlations across the series. The null and alternative hypotheses are as follows:

$$H_0: \delta_i = \delta_2 = \cdots \ldots = \delta_n = \delta = 0 \qquad (4.37)$$

H_1: δ_i<0 for at least one individual (i) in the series (4.38)

For the test to be accurate, the panel does not have to be balanced. In other words, the time series component of the panel does not need to be the same for all the cross-sectional units. The t-statistic in the test is computed by using the average value from the individual ADF t-statistics. That is, testing for δ_i=0 in all i

$$\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_{\delta_i} \qquad (4.39)$$

t_{δ_i} represents each individual ADF t-statistic using the null hypothesis. Im *et al.* (2003) assumed that t_{δ_i} has a finite mean and variance if it is independent and identically distributed (IDD), then as T$\rightarrow \infty$ we have

$$\frac{\sqrt{N}(\frac{1}{N}\sum_{i=1}^{N} t_{iT} - \frac{1}{N}\sum_{i=1}^{N} E[t_{iT}/\delta_i = 0])}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} var[t_{iT}/\delta_i = 0]}} \Rightarrow N(0,1) \qquad (4.40)$$

Likewise as N$\rightarrow \infty$, using the Lindeberg-Levy central limit theorem,

$$t_{lps} = \frac{\sqrt{N}(\bar{t} - N^{-1}\sum_{i=1}^{N} E[t_{iT}/\delta_i = 0])}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} var[t_{iT}/\delta_i = 0]}} \Rightarrow N(0,1) \qquad (4.41)$$

Both the values of $E[t_{iT}/\delta_i = 0]$ and $var[t_{iT}\ /\delta_i = 0\ ]$ were computed using stochastic simulation and tabulated in their paper by the authors.

4.4.6 Other panel unit root tests

Maddala and Wu (1999) and Choi (2001) proposed a Fisher-type test which uses *p*-values based on the Dickey Fuller's approach, to validate the null of a unit root. The test is defined as:

$$P = -2\sum_{i=1}^{N} ln_{\delta_i} \qquad (4.42)$$

-$2ln_{\delta_i}$ has two degrees of freedom and it comes from a x^2 distribution. For finite N, as $T_i \rightarrow \infty$, P has 2N degrees of freedom and it is distributed as x^2. Maddala and Wu (1999) pointed out that the test proposed by Im *et al.* (2003) and their Fisher type eliminates the assumption made by Levin *et al.* (2002) that δ_i is identical across the cross-sectional units.

The test has many advantages over both LLC and IPS because of its ability to work well even when the panel is unbalanced, easy computation and it can deal better with cross-sectional dependence across the series. For more units root tests for panels, see for example Breitung (2000), and Karlsson and Lothgren (2000).

Breitung (2000) used the Monte Carlo experiment to obtain a test statistic, that excludes a bias adjustment and is better than both the LLC and IPS in terms of its ability to work well even when either N is small or large in comparism with T.

Karlsson and Lothgren (2000) evaluated the power of both LLC and IPS tests using different panel sizes. They found that there is a risk of concluding wrongly that the whole panel is stationary when the times series is long, and likewise when the time series is short, there is a risk of concluding the panel has unit roots even when most of the series are stationary. The authors proposed that in order to know if the panel is stationary or not, the researcher should carry out both the individual and panel unit roots test.

Structural breaks were considered, for example, in the unit root tests developed by Culver and Papell (1997), and Murray and Papell (2000). Also, Pesaran (2004) explored how to deal with cross-sectional dependence when testing for unit roots in a panel data context. It is evident from the discussion in sections (4.4.5) and (4.4.6) that econometric analysis requires that when using macroeconomic data that span over a long period of time, the variables must be tested for unit roots. In other words, as mentioned earlier, to avoid the possibility of spurious results in the series, the series must be tested to ensure that they are stationary in levels and/or after differencing them. Or at least, the order of integration of the variables should be known by the researcher. Hence, assessing the unit root properties of the variables in a time series data is an important requirement in pure time series analysis as well in panel data analysis.

The use of panel data improves the efficiency of the unit root test, including the Augmented Dickey Fuller's approach, as argued by Levin, Lin and Chu (2002) because of the large sample size in a panel unit. Therefore, since evidence has shown that time series data possess memory of the past and panel data is a combination of both cross sectional data and time series data, it is important to establish if the series are stationary or not. It follows that when testing for cointegration in panel data, undertaking unit root tests is mandatory. The concept of cointegration is explored in the section that follows.

4.4.7 Cointegration tests in panel data

The idea of testing for cointegration in panel series is analogous to the context of time-series data, as can be gauged from the discussion of unit root tests above. In essence, cointegration tests aim at establishing whether the variables in a model move together in the long run. As discussed in section (4.4.1) above, a regression is said to be spurious when we conclude wrongly from the results that a relationship exists between a dependent variable (Y) and one or more independent variables.

For studies on spurious regression in panel data see, for example, Entorf (1997), Phillips and Moon (1999) and Sun (2004). In the next sub-sections, some of the panel cointegration tests in the literature are explained, and the chosen method for the aggregated estimation in this PhD study is also highlighted in the discussion. Another difference between the two types of regressions is the possibility of a linear combination of the dependent and first difference of the independent variables being stationary. As explained by Harris (1995, p. 22):

''*the economic interpretation of cointegration is that if two or more variables are linked to establish an equilibrium relationship spanning the long run, then even though the variables have a unit root, they will eventually move together over time due to the mean reverting behaviour in time series data, and the difference between them will be stable''*.

This applies in the context of panel data employed in this research study, because it has time series (T=34) element which suggest that if the variables are cointegrated, there is evidence of a long run equilibrum in which the variables converge to over time. Baltagi and Kao (2000) also assert that, panel cointegration frameworks are used to answer questions about the long run economic relationships found in macroeconomic and financial data. The authors stated further that economic theory is used to predict if such relationship exist in the series, by interpreting the regression coefficients.

4.4.8 Residual based tests for panel cointegration

The residual based tests proposed by Kao (1999) based on Dickey-Fuller (DF) and ADF for homogeneous panels are discussed below. Other authors proposing a residual based tests include: McCoskey and Kao (1999), Phillips and Moon (1999) and Pedroni (2004).

From the panel regression model

$$y_{it} = x_{it}^{l} + v_{it}^{l} + u_{it} \qquad (4.43)$$

In (4.43), x_{it} and y_{it} are stationary in first difference (I(1)) and not cointegrated, $v_{it} = \{\mho_i\}$. Kao (1999) argued that the DF and ADF approaches should be used to estimate the residuals, using a null hypothesis of no cointegration, H_0: $\psi = 1$. The estimated residuals from the DF test of equation (4.43) can be computed from

$$\hat{u}_{it} = \psi\hat{u}_{it-1} + \pi_{it} \qquad (4.44)$$

$\hat{u}_{it}$ stands for the estimated residuals and π_{it} is the white noise disturbance term. The OLS estimate of ψ, and the t statistic can be expressed as:

$$\hat{\psi} = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{u}_{it}\hat{u}_{it-1}}{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{u}_{it}^{2}} \qquad (4.45)$$

$$t_{\psi} = \frac{(\hat{\psi}-1)\sqrt{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{u}_{it-1}^{2}}}{s_u} \qquad (4.46)$$

where $s_u^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=2}^{T}(\hat{u}_{it} - \hat{\psi}\hat{u}_{it-1})^2$. Kao (1999) introduced the five DF type tests shown below

I. $DF_{\psi} = \frac{T\sqrt{N}(\hat{\psi}-1)+3\sqrt{N}}{\sqrt{10.2}}$

II. $DF_t = \sqrt{1.25}t_{\psi} + \sqrt{1.875N}$

III. $DF_{\psi}^{*} = \frac{\sqrt{N}T(\hat{\psi}-1)+\frac{3\sqrt{N}\hat{\sigma}_{v}^{2}}{\hat{\sigma}_{0v}^{2}}}{\sqrt{3+\frac{7.2\hat{\sigma}_{v}^{4}}{\hat{\sigma}_{0v}^{4}}}}$

IV. $DF_{t}^{*} = \frac{t_{\psi}+\frac{\sqrt{6N}\hat{\sigma}_{v}}{2\hat{\sigma}_{0v}}}{\sqrt{\frac{\hat{\sigma}_{0v}^{2}}{2\hat{\sigma}_{v}^{2}}+\frac{3\hat{\sigma}_{v}^{2}}{10\hat{\sigma}_{0v}^{2}}}}$

V. $ADF = \frac{t_{ADF}+\frac{\sqrt{6N}\hat{\sigma}_{v}}{2\hat{\sigma}_{0v}}}{\sqrt{\frac{\hat{\sigma}_{0v}^{2}}{2\hat{\sigma}_{v}^{2}}+\frac{3\hat{\sigma}_{v}^{2}}{10\hat{\sigma}_{0v}^{2}}}}$

t_{ADF} stands for the t-statistic of ψ in the ADF regression

$\hat{u}_{it} = \psi\hat{u}_{it-1} + \sum_{j=1}^{p}\phi_j\Delta\hat{u}_{it-j} + \pi_{it}$ (4.47)

The asymptotic distribution of $DF_{\psi},\ DF_{t},\ DF_{\psi}^{*},\ DF_{t}^{*},\ and\ ADF$ converge by sequence limit theory to a standard normal distribution N (0,1).

The null hypothesis of cointegration was employed in the residual based test proposed by McCoskey and Kao (1998). Rooted in the time series literature, the authors used the Langrange Multiplier (LM) test and the Locally Based Invariant (LBI) test to derive the model for a panel unit. The model proposed, which allows for variation in the slopes and intercepts, can be written as:

$y_{it} = \delta_i + x_{it}^{\iota}\beta_i + u_{it}$ (4.48)

$x_{it} = x_{it-1} + \epsilon_{it}$ (4.49)

$u_{it} = \xi_{it} + v_{it}$ (4.50)

$\xi_{it} = \xi_{it-1} + \theta v_{it}$ (4.51)

where v_{it} are independently and identically distributed (IID) (0, σ_p^2), H_0: $\theta = 0$, that is cointegration. The t-statistic proposed by McCoskey and Kao (1998) is:

LM = $\frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}z_{it}^2}{\hat{\sigma}_e^2}$ (4.52)

where z_{it} is the partial sum process of the error terms and is defined as, $z_{it} = \sum_{j=1}^{t}\hat{u}_{ij}$, and $\hat{\sigma}_e^2$ is defined by the authors. The test asymptotic result is

$$\sqrt{N}\left(LM - \mu_q\right) \Rightarrow N\ \left(0, \sigma_q^2\right) \tag{4.53}$$

Monte Carlo simulations can be used to find the value of μ_q and σ_q^2. One of the merits of using the LM approach lies in its ability to eliminate heteroscedasticity in the estimated parameters. Baltagi (2008) argues that the test would have been more robust if it were capable of accounting for cross-sectional dependence. The author also pointed out that the asymptotic theory which underpins the method, does not give a good approximation when employed in empirical research.

4.4.9 Westerlund (2007) error correction panel cointegration test

To determine the existence or otherwise of a long run (cointegrating) relationship in a model, the Westerlund (2007) panel cointegration test can be used. The test was proposed by Persyn and Westerlund (2008), the DGP used by the error correction model is

$$\Delta y_{it} = \lambda_1^1 d_t + \alpha_i\left(y_{i,t-1} - \beta_i^1 Z_{i,t-1}\right) + \sum_{j=1}^{p_1} \alpha_{ij}\, \Delta y_{i,t-j} + \sum_{j=-q_i}^{p_i} \gamma_{ij}\, \Delta Z_{i,t-j} + u_{it} \tag{4.54}$$

where t=1,………..,T, for the time series

i=1,………...,N, for the cross sectional units

d_t contains the deterministic components, it has three cases

in case 1, d_t =0 and it has a deterministic term

in case 2, d_t =1, Δy_{it} is generated with a constant and

case 3, d_t = (I, t)^1 so that Δy_{it} is generated with a constant and a trend.

Also, ΔZ_{it} is independent of the error term u_{it}, and t and I are also independent of the error term.

Persyn and Westerlund (2008) assert that any dependence across the cross-sectional unit (i) can be analysed using the bootstrap technique to improve the performance of the test.

Persyn and Westerlund (2008) present the model used as:

$$\Delta y_{it} = \psi_1^1 d_t + \alpha_i y_{i,t-1} + \lambda_t^i Z_{i,t-1} + \sum_{j=1}^{p_1} \alpha_{ij}\ \Delta_{ij,t-j} + \sum_{j=-q_i}^{p_i} \gamma_{ij}\, \Delta Z_{i,t-j} + u_{it} \qquad (4.55)$$

where λ_t^i=- $\alpha_i\, \beta_i^1$, and the rate at which the equilibrium position $y_{i,t-1} - \beta_{i,t-1}$ is adjusted back to its equilibrium level after a shock is determined by α_i.
Therefore, the panel error correction cointegration test by Westerlund (2007) is used to test the null hypothesis of no cointegration, which is investigated by testing if any member of the panel is 'error correcting' or not. That is: H0: α_i =0. The decision criterion is based on whether the error correction term is equal to zero. The test is robust enough to allow for a high level of heterogeneity in the short run as well as in the long run cointegrating relationship, even in an unbalanced panel.

The Westerlund (2007) error correction panel cointegration test provides four test statistics that is Gt, Ga, Pt and Pa. The Gt and Ga test the null hypothesis that H0: a_i= 0 for all i against the alternative hypothesis that H1: a_i < 0 for at least one of the variables in i, which is a weighted average of each of the variables in i and their corresponding t-ratios. On the other hand, Pt and Pa test statistics analyse the panel cross sectional units using the pooled information to test the null hypothesis that H0: a_i = 0 against the alternative hypothesis that H1: a_i <0 for all the cross sectional units. Therefore, the rejection of the null hypotheses (H0) would mean that there is cointegration in the panel. The rejection of the null hypothesis could stem from a rejection from any of the cross sectional unit of the panel, and it is not possible to know from the test which of the cross section unit led to the rejection of the null hypothesis.

4.5 Pooled mean group (PMG) estimator error correction test

Pesaran, Shin and Smith (1999) proposed a maximum likelihood (ML) pooled mean group (PMG) estimator, for heterogeneous dynamic panels. The method facilitates economic interpretation by specifying an error correction equation which fits an Autoregressive Distributed Lag (ARDL) model into the data (Alagidede *et al.,* 2014). The error correction model of ARDL (p,q,q,……q) proposed by Alagidede *et al.* (2014), can be written as:

$$y_{it} = \Phi_i y_{i,t-1} + \beta^1 Z_{it-1} + \sum_{j=1}^{p-1} \lambda_{ij} \Delta y_{i,t-j} + \sum_{j=0}^{q-1} \lambda_{ij} \Delta Z_{i,t-j} + \mu_i + \varepsilon_{it} \qquad (4.56)$$

where Z- vector of explanatory variables

β_i- consists of the information about the long run impacts

ϕ_i- error correction term

λ_{ij}- Includes the short run properties or information.

The Pooled Mean Group (PMG) estimator is an intermediate between the Mean Group (MG) estimator and the fixed and random effects estimator. The Mean Group estimator takes the average of the variables, while assuming homogeneity across countries. On the other hand, the fixed and random effects estimator only allows the intercept to vary, other coefficients are assumed to be the same across the group (Pesaran, Shin and Smith, 1999). However, PMG uses pooling, which involves long run homogeneity across the countries and averaging across the group, and is used to analyse the short run parameters in the model as well as the error correction coefficient (Pesaran, Shin and Smith, 1999). It also allows the intercept, short run coefficients and error variances to be different across the groups. But the long run coefficient is assumed to be the same for all the groups, that is $\beta = \beta_i$ for all I (Pesaran, Shin and Smith 1999). Pesaran *et al.* (1999) argued further that since there are similarities like technology, arbitrage conditions and solvency constraints across the countries, the variables in each country will have identical long run equilibrium. This is true for the sub Saharan African countries, as they all have a wide energy demand to supply deficit and use similar technology.

Hence, this present study will analyse the short run and long run impact of income, price of energy and other identified variables from the literature on the aggregate demand for energy in the Sub-Saharan African countries using the pooled mean group estimator (PMG) error correction method.

4.6 Other panel cointegration tests

Pedroni (2000, 2004) developed a test that allows for heterogeneity in the cross-sectional units using the null hypothesis of cointegration. The test which is in twofold, namely, one that tests for cointegration by taking the average across cross-sectional units, and one that takes the average of both the numerator and the denominator. The former, takes the form of:

$$\hat{Z}_{\psi} = \sum_{i=1}^{N} \frac{\sum_{t=1}^{T}(\hat{u}_{it-1}\Delta\hat{u}_{it}-\hat{\lambda}_i}{(\sum_{t=1}^{T}\hat{u}_{it-1}^2)} \qquad (4.57)$$

where $\hat{u}_{it}$ is computed from equation (4.34), $\hat{\lambda}_i = 2^{-1}(\hat{\sigma}_i^2 - \hat{s}_i^2)$, and the individual long-run and contemporaneous variances of the error term are represented by $\hat{\sigma}_i^2$ and $\hat{s}_i^2$ respectively.

For the 2004 paper, the author uses four panel variance ratio statistics, which are presented in his paper.

Larsson *et al.* (2001) used average individual rank trace statistics to develop a likelihood based panel cointegration test for heterogeneous panel units. The likelihood-based approach was also used by Groen and Kleibergen (2003) to present a method for testing for cointegration in a fixed number vector error correction model (VECM). The method by Groen and Kleibergen (2003) allows for cross-sectional correlation and can be used for both homogeneous and heterogeneous cointegrating vectors.

4.7 Chapter summary

This chapter detailed the main methodological framework adopted in this PhD study. The specification and advantages of using the panel data framework were discussed, before moving on to discuss the main classifications of panel data models. In the first category that explored static linear models, fixed effects and random effects models were explained in turn, and what could be done to select the appropriate model for analysis. After examining this, it was pointed out that if one of the properties of a good regression model is violated then the Prais-Winsten regression can be used to obtain more reliable estimates. This is in line with the econometrics literature cited earlier (see Section 4.3.5). These models will be used to analyse the disaggregated energy demand in SSA because the dataset is fairly balanced.

The other model category discussed was non-stationary panel models. The basic concepts of stationarity and unit roots were explained based on the time series analysis, as a background to appreciate the greater complexities that pertain to the application of these concepts in the context of panel data models. The pros and cons of the various approaches used in the empirical literature were highlighted. The issues related to the use of a heterogeneous panel and cross-sectional dependence were also examined. Based on the merits and suitability of IPS, ADF-Fisher and PP- Fisher unit root tests, Westerlund (2007) cointegration tests, and the Pesaran *et al.* (1999) PMG estimator will be used for the econometric analysis of the aggregate demand for energy in SSA presented in the next chapter, as they are particularly appropriate in the case of an unbalanced panel such as the one used in the present study.

The data employed for the study and the measurement issues are examined in the next chapter.

Chapter 5 : Data Description and Sources

5.1 Chapter overview

This chapter discusses the dataset employed for the econometric analysis presented in chapter 6. In order to construct reliable econometric energy models and obtain valid results from which relevant inferences about energy policies can be drawn, it is important to have consistent long time series data on energy demand and the factors that influence it (Pesaran *et al.,*1998). Therefore, in this part of the thesis each of the variables used are described for both the aggregate and the disaggregated energy demand in Sub-Saharan Africa (SSA). The sources of each of the variables used in the dataset are also given, and the measurement problems associated with them is discussed.

The chapter also elaborates on some of the benefits and drawbacks of the use of secondary data in econometric analysis. The chapter concludes by justifying the use of the secondary data employed in this study through the data cleaning procedure employed.

5.2 Content of the panel dataset

The panel dataset used for the regression analysis contains annual observations for the period 1980 to 2014, the most up-to-date data that were available during the data collection phase of this study. Due to some data constraints and limited data availability in most of the countries in the region, the dataset constructed includes data for 16 countries with few missing observations in Sub-Saharan Africa. However, Cote D'Ivoire, Ghana, Mozambique and Zimbabwe are not analysed in the disaggregated analysis by fuel type due to a high number of missing values for most energy types. The ones included are the countries in the region with most complete available data from the International Energy Agency (IEA) database for non-OECD countries energy summary, for the period between 1980 and 2013 (see Table 5.1 below for a full description of the list of countries and the associated sub-region).

Table 5.1: Table showing the countries included in SSA panel dataset

Sub-region	Country
Southern Africa	Botswana (BWA)
	Mozambique (MOZ)
	South Africa (ZAF)
	Zambia (ZMB)
	Zimbabwe (ZWE)
Central Africa	Cameroon (CMR)
	Congo (COG)
	Democratic Republic of Congo (COD)
West Africa	Benin (BEN)
	Cote D'Ivoire (CIG)
	Ghana(GHA)
	Nigeria (NGA)
	Senegal (SEN)
	Togo (TGO)
East Africa	Ethiopia (ETH)
	Sudan (SUD)

Source: Researcher's compilation (2015)

5.3 Data organisation

The data are collected for each of the following variable categories:

- Energy consumption
- Other economic data
- Energy prices

Each of the categories listed above is discussed one after the other in the sections below, and where applicable the number of missing observations for any of the data series is also highlighted.

5.3.1 Energy consumption

The energy consumption data is disaggregated by six relevant liquid fuel types (electricity, gasoline, diesel, LPG, solid biomass and kerosene), using measures that are typically employed in relevant literature. The liquid fuels are measured in thousand Tons of oil Equivalent (TOE). Electricity is measured in gigawatt hours while biomass is measured in Tera joules. Consumption data for electricity is also included alongside the consumption data for biomass which is used traditionally by the majority of the consumers in the region. Motor gasoline refers to the liquid fuel used in internal combustion engines like vehicles, construction equipment, trucks and alternative power generating sets measured in kilotonne (IEA, 2015). The final consumption data contains the total of the energy used by all the sectors in the economy. For instance, the total motor gasoline used by the Industry, Residential, Transport, Industrial, Agriculture and the public sector of the economy. Liquefied Petroleum Gas (LPG) is defined by the IEA (IEA, 2015) as the light hydrocarbon component of the paraffin family which contains either propane or butane, or the combination of both gases. Gas/diesel oil is heavy gas oils used in different sectors of the economy.

Solid biomass included in the dataset represents plants used directly or indirectly as fuel before burning. It includes different types of woody products from the forest, agricultural or industrial processes obtained from firewood, sawdust, and wood shavings. It is measured in Tera joules based on a net calorific basis (IEA, 2015).

Energy consumption data for Botswana starts from 1981. Before this period it was classified under 'other Africa' in the International Energy Agency (IEA) database. The data for Eritrea was included under Ethiopia before 1992, which accounts for the missing observations between 1980 and 1991 according to IEA.

Aggregate energy use includes the primary form of energy before it is transformed to other end-use fuel which is the sum of local production plus imports and stock changes. It excludes the amount exported or used by ships and aircraft employed in international transport (IEA statistics as citied in WDI, 2016).

5.3.2 Other economic data

The category 'other economic data' includes other economic factors that impact on the demand for energy. The definition of each variable used in the study is provided in this section of the chapter.

Gross Domestic Product (GDP) at international price (constant, 2005, in US$) is used for all the countries. According to the definition provided in the World Bank database, GDP per capita is gross domestic product divided by midyear population. GDP is the total addition of the value added by all resident producers in a country plus the inclusion of all taxes on products and the deduction of all subsidies not included in product's value (WDI, 2015). For each of the countries, the World Bank uses the 2005 official US dollar exchange rate to convert the GDP from domestic currencies to the international figures, as at when the data for this research was obtained. For both the aggregate and disaggregated energy demand models, GDP per capita figures are used.

The 'total population' data category is used as a measure for the population variable. The total population for a country includes all the people who are resident of the country with no consideration given to their citizenship status. However, it excludes the number of refugees who are dwelling temporarily in the country of asylum (WDI, 2015). All the time series data for population are consistent and complete for all countries in the study, as there are no missing observations.

Industry value added is the figure of the value added by the mining, manufacturing, construction, electricity, water and gas sectors in an economy. On the other hand, services value added is the value added by both the wholesale and retail categories of hotels, restaurants, transport, government, financial, educational, healthcare and the real estate sectors. Value added is defined as the net output of a sector after adding up all outputs and subtracting the figures for intermediate inputs (WDI, 2015). The data for these variables have some missing observations.

The economic structure of a country can be obtained by dividing the real value added in industry by that of the real value added in service sectors (see, for example, Mensah *et al.,* 2016). This will be used as a proxy for the economic structure of each of the countries in the study.

The 'annual percentage growth rate of urban population' from WDI is used as a measure of the degree of urbanisation. The people living in the urban areas of a country are known as the urban population. All the time series data for the urban population growth rate are consistent and complete for all countries in the study, as there are no missing observations. The variable is measured in percentage change.

Due to limited data availability for the region, energy prices are not available for most of the countries in Sub-Saharan Africa. Even when energy prices series are available, they are very limited and incomplete. The average annual price for crude oil is measured in US dollars per barrel, and the data are complete for the period of the study, that is from 1980 to 2014.

5.4 Sources of data

The disaggregated data for energy consumption were obtained from the International Energy Agency (IEA) database via the UK Data service website. The data were extracted and imported into an excel file, which was then used as part of the Sub-Saharan Africa dataset compiled for the study. The aggregate energy data was obtained from the World Bank Development Indicator database.

The economic data (industry value added, services value added, GDP, total population and urban population growth) were obtained from the World Bank database, known as the World Development Indicators (WDI). The data were downloaded into excel, before combining them with other data in the SSA dataset. Crude oil prices are taken from the International Monetary Fund (IMF) database under the International Financial Statistics category. It was accessed via the UK data service, from which it was downloaded into Microsoft excel. It is in US Dollars.

5.5 Measurement problems

Secondary data refers to data that has already been collected (Byrne, 2002). Before discussing the measurement issues which may arise as a result of the use of secondary data, cross-country analysis or the variables used, some of the advantages of using secondary data are first given below:

- As most of the data used in energy and economic modelling are of interest to both international and national stakeholders, funded studies and international agencies have collected and continue to collect large and quality data over several years. These data can be accessed and used to answer new questions, which can provide new insights to several energy issues.
- As with any statistical study, the larger the sample size the more reliable the inferences that can be drawn about the population. With the use of secondary data, large data is available which can be used to provide more robust findings due to a higher degree of freedom (Smith, 2008 as cited in Johnson, 2014). With the long time series data over the 33 years period used in this study, the cross-country model results will certainly be more robust.
- The use of secondary data enables the use of several methods, models, and a new interpretation of previously used data through the application of new ideas and models to give new perspectives. In this study, existing data are used to construct the Sub-Saharan Africa model which will provide new knowledge on the energy demand debate for the region.

Despite the merits of the use of secondary data, it has a few drawbacks which were carefully considered in this study during the data collection process. One of the drawbacks relates to the problem that the original data had been collected for a different objective. For instance, the data could have been collected at a different region or country (Johnson, 2014).

This study does not suffer from this drawback because the secondary data used were collected in the country of interest and can be purposefully used for our aim. Data source was verified by reading and studying each data source database documentation where all the definitions, measurement units and each country profile that details the data collection methods are discussed (for example, see the World Energy Statistics 2015 revised edition by the IEA on how each country energy data was collected by the agency).

From the discussion above it seems evident that the use of secondary data is the best approach to answer the quantitative research questions in this PhD study, especially considering the geographical location and the number of countries covered in the present study. Despite the advantages of the use of secondary data, there are still some measurement issues peculiar to the dataset. Some of these issues are discussed in the sub-sections below.

5.5.1 Energy use measures

The measure of energy consumption considered in this PhD study includes energy derived from traditional biomass. This is because the countries studied produce a substantial amount of energy from this source. In other words, the total final consumption of both modern energy (gasoline, diesel, kerosene, LPG and electricity) and traditional sources (biomass) are analysed.

This seems to be the best approach as it will address some of the shortcomings of previous studies where biomass is not included in the model (for literature on this, see Chapter 3). From the data, it is evident that solid biomass source like firewood is used vastly in the residential sector. Biomass consumption in the residential sector made up more than 70% of the total biomass consumed in most of the countries analysed in this study (see appendix 2). This is in line with the existing literature and publications about the SSA energy mix (e.g. Lambe *et al*., 2015). Therefore, the energy consumed from biomass cannot be ignored in any rigorous econometric study of the energy demand in SSA.

The IEA database provides all the disaggregated energy consumption data using the same measurement for all countries. This is used in this study to aid data analysis, comparability and interpretation.

5.5.2 Output measures

For GDP data, the pricing system is not the same in all the countries and the amount contributed by primary commodities also differs. The best approach, therefore, is to use data in international prices in US dollars. Even though some of the individual countries have a database with their economic statistics, there was no point in getting data from each individual country database as the World Bank database contains compiled data by national statistical offices and provides them in a user friendly and accessible platform to the public. Moreover, for consistency of measurement of the series, it is also risky to mix measures of the same variable for different countries from separate databases.

However, the GDP data reported in US dollars at 2005 constant prices in some countries were verified. For instance, by checking the Nigerian statistical bulletin (NSB, 2014), the study was able to compare the GDP data in WDI with the data reported in the bulletin. In both sources, data figures were very close with only minor differences, probably due to rounding up.

The World Bank also states in the Database Handbook that they employ the use of General Data Dissemination Service (GDDS) developed by The World Bank and the IMF. This is used to ensure that the data reported by each country statistical offices are reliable, consistent and follow international guidelines. Through the system the bank is also able to provide frameworks and guidelines, which the national statistical offices of participating countries can adopt to improve their knowledge of how to collect, and distribute comprehensive and timely data. On this basis, I have used the GDP data in international prices as provided by the World Bank due to their quality, reliability and comparability.

5.5.3 Other economic data measures

According to the WDI manual, population data were taken from the United Nations World population prospects. Such data were verified in the study by making sure the reported trend corresponds to the expected growth pattern, and in some cases other databases were checked to ensure consistency.

The economic structure of most countries has shifted from manufacturing to more service oriented sectors (Mensah *et al.,* 2016); the research will shed more light as to how this has impacted on the demand for energy in SSA. International prices are used to enable cross-country analysis.

Urban population growth rate data seem to have the most measurement concern due to the lack of a uniform definition of urban area globally (WDI, 2015). According to the WDI manual as there is no universal way of differentiating between rural and urban centre, it may be difficult to measure this adequately across countries. However, the figures are computed using the World Bank population estimates and the United Nations World Urbanisation prospects ratio, and extracted from the same database for all the countries under study, the data is expected to be reliable as the same models were used to arrive at the values.

5.6 Data cleaning

Data cleaning which involves how missing observations, outliers and inconsistencies in variable flows, among others, are considered in a study, have been the subject of much academic debate in the applied literature. Some authors have argued that data cleaning carried out by the researcher should be properly documented and reports should be provided in empirical studies (Broeck *et al.,* 2005; Dong and Peng, 2013; Young and Johnson, 2015). Failure to do so may result in biased coefficients of parameters (Dong and Peng, 2013). Therefore, in this section the issue of data cleaning is discussed alongside the assumptions made with regard to some of the potential concerns. The mechamism of missing data is first discussed and how it relates to the SSA dataset, the implications of the proportion of missing data and how the dataset was cleaned is provided below.

Rubin (1976) argues that there are three main mechanisms of missing data: missing completely at random (MCAR); missing at random (MAR); and missing not at random (MNAR). To illustrate each category of missing data, we divide a complete dataset represented by X into two parts: X_{obs} and $X_{mis.}$ Where X_{obs} is the observed part in X, while X_{mis} is X missing part. For the MAR case according to Rubin (1976), the probablility of missingness only depends on the observed and not the missing data. Whereas in MCAR case which is a special case of MAR (Schafer and Graham, 2002), the probablilty of misssingness does not depend on any part of X (that is, neither the observed nor missing part). In the last category MNAR, the probablity of missingness depends on the missing data. From the illustration, it could be inferred that missing data that follow under the MAR category may be ignored but those under MNAR should not be ignored (Schafer and Graham, 2002).

There seem to be a consensus among the studies who explored the impact of missing values in empirical research that if 'missingness' is not under the researcher's influence and the distribution of the missing data is unknown, MAR is the best assumption (Schafer and Graham, 2002; Dong and Peng, 2013). This is corroborated by Collins *et al.* (2001) who argue that parameter estimates and standard errors are only slightly biased even when the assumption of MAR is violated.

Based on the above, assuming that the missing observations in the SSA dataset is MAR seems both plausible and reasonable considering that the missing data appear not to depend on the observed data distribution. Besides, in the case of the disaggregated energy consumption data, IEA stated why some of the data for the countries were missing. The reasons given had no link to the observed data because some were due to the period in which the countries were still colonised, which is clearly out of researcher's control. Moreover, since the proportion of missing observations in the SSA dataset is about 5% of the total, the estimates are not likely to be biased. This is supported by Schafer (1999), who found that the missing rate of 5% in a dataset may not have any significant impact on the estimates. As one would expect, the mechanism behind the missing data will have more impact on the empirical results than the proportion of the missing observations (Tabachnick and Fidell, 2012). Therefore, if a large number of observations is used in a study the percentage of missing variables should not lead to bias in the estimates.

The approaches used in the literature to handle missing data include: the pairwise deletion method; listwise deletion method; multiple-imputation; full information maximum likelihood; and the expectation-maximisation method. The last three which are more recent are referred to as principled methods by researchers (see, for example, Dong and Peng, 2013). Of all the above-mentioned methods, listwise deletion appears to be the most commonly and widely used approach in the literature (Breitwieser and Wick, 2015). An example given in the study by Breitwieser and Wick (2015) reports that in a panel data context, using the listwise deletion approach would mean that all countries with missing data in some years will be automatically excluded in the analysis by the statistical software used. This was exemplified by Young and Johnson (2015) who asserted that the xt and st commands used in STATA (the statistical software used in this study to perform the regression analysis) will still work well even if there are missing observations under some variable series in a panel. The xt procedure will be used in the data analysis of the SSA dataset presented in the next chapter.

Having addressed the issue of missing observations, other data cleaning procedures carried out on the SSA dataset are discussed. Firstly, all data were correctly entered into the excel workbook to eliminate entry error. During the dataset entries, the pattern for each variable in individual countries, was checked for any irregular trend and if spotted, two different databases were checked for the same data to make sure that they were similar to the data from the database used. For instance, the population and GDP data, were checked in the WDI, IMF and OECD databases to ensure consistency and accuracy of data. All of them had the same pattern with very little differences in the figures which could be due to rounding up.

Secondly, all data were checked for outliers or inconsistencies. To do this, individual country data for some variables were graphed so as to get a good feel for the data and thus ensure that there is no strange or anomalous pattern. During this process, missing observations were revisited again to check if there are specific reasons for the omission. However, for cases such as energy prices, there was no explanation other than the fact that there are limited and inconsistent energy price data for Sub-Saharan African countries. This is one of the limitations of the dataset of the present study.

Thirdly, the summary statistics of the SSA dataset were presented in a tabulated form to check properties of the data (see Chapter 6). Lastly, the natural logarithms of the variables are taken so as to use a log-linear model for the analysis. This will ensure that estimated coefficients can be interpreted as elasticities, and help in interpreting results.

Furthermore, the use of a panel data technique which works well under missing observations is important when working with an unbalanced panel dataset like the SSA dataset used in this study. Therefore, the use of a panel data technique with the mentioned strength seems to be the best approach for achieving a good match between the research question and the data available. The pooled mean group (PMG) estimator (see Chapter 4 on panel data estimators) can handle missing observations and will be used for the aggregate analysis reported in Chapter 6. There is no issue of missing values in the disaggregate dataset, therefore panel linear models are employed for the analysis also presented in the next chapter.

5.7 Chapter summary

The purpose of this chapter was to give a detailed description of the SSA panel dataset used for the analysis of energy demand in Sub-Saharan Africa presented in the next chapter. This was done by explaining the definitions of the variables used, sources of the data, measurements used and the data cleaning carried out on the dataset. A critical discussion of how data measurement issues were handled, for example, in relation to missing observations, was also provided. From the discussion offered in this chapter, it is evident that the research was able to collect and make use of reliable, high quality data for the analysis by using international and reputable energy and economic databases.

In Chapter 6, the econometric analysis is presented, and the results of the estimations properly interpreted.

Chapter 6 : Econometric Analysis and Estimation Results

6.1 Chapter overview

This chapter presents the results of the estimations (obtained using the econometric software STATA 13), following the econometric strategy developed to analyse the determinants of energy demand in Sub-Saharan Africa (SSA) as discussed in the methodology chapter (Chapter 4). The econometric procedure essentially can be divided into three stages. First, the descriptive statistics of the dataset used will be examined alongside the analysis of the properties of each variable included in the regression through visual inspection of the plots of the relevant series and formal unit root tests. The *a priori* expectations of the relationship between the dependent variable and each of the independent variables will also be stated. Second, the cointegration tests will be presented and the significance of the long-run estimated coefficients discussed. Third, the results of the corresponding error correction model (ECM) will be shown and discussed.

This sequence of analysis will be undertaken for the aggregate demand relationship, and the linear panel models in the context of fixed effects, random effects and PW models used in balanced panel models presented for the disaggregated analysis. The results of the econometric analysis disaggregated by energy types is based on the determinants of demand for kerosene, petrol (gasoline), liquefied petroleum gas (LPG), biomass, diesel and electricity in SSA. The last section of the chapter concludes.

6.2 APriori expectations of the estimated aggregate relationship

Prior to the presentation of the aggregate and disaggregated results, it bears reminding what - according to theory - are the *a priori* expectations in terms of the sign of the estimated parameters and hence the expected relationship between the variables analysed in the present study.

H1: Positive income elasticity is expected because as the income level of consumers' increases, they increase the amount of commodities in their consumption bundle, which includes energy. This assumption is rooted in the demand theory discussed in Chapter Two.

H2: Also in line with demand theory, a negative price elasticity is expected because theory suggests that as the price to be paid for commodities increases, consumers will reduce their consumption level. Thus, an increase in energy price is expected to reduce the amount of energy consumed.

H3: Growth in urban population known as 'urbanisation' is expected to have a positive impact on the aggregate demand for energy. This is because as people move from rural to urban areas, it is assumed that they change from the traditional forms of energy like solid biomass to modern energy types like electricity and LPG.

H4: A positive relationship is expected between energy demand and economic structure. The economic structure variable is derived by dividing the share of industrial value added to the services value added, which is then used to measure the impact of structural changes in the countries. Therefore, as a country increases its industrial output, more energy is used, which increases the total energy consumed in the country.

H5: An increase in population will increase the amount of energy used in SSA by consumers. Therefore, a positive relationship between energy demand and population is expected.

6.2.1 Descriptive statistics of the variables

Table 6.1 presents the basic characteristics of the data used to analyse the driving forces of energy demand in SSA. In other words, it gives a simple summary of the dataset used for the analysis in this chapter. For instance, from the table we can see that the total number of observations is 2678 from the sixteen countries analysed in the 33 years making up the sample period.

From the table, it is evident that we have an unbalanced panel which also informed the selection of the model used for the analysis as discussed in Chapter 4.

Table 6.1: Summary statistics of the SSA panel dataset

Variable		Mean	Std Dev	Min	Max	Obs
lnDemand	Overall	6.08	0.66	5.25	8.00	526
	Between		0.64	5.25	7.88	
	Within		0.22	5.53	6.78	
lnIncome	Overall	6.55	0.95	4.74	8.95	559
	Between		0.94	5.07	8.57	
	Within		0.27	5.64	7.27	
lnOilprice	Overall	3.43	0.66	2.57	4.65	560
	Between		0	3.43	3.43	
	Within		0.66	2.57	4.65	
lnUrban	Overall	1.35	0.40	-1.30	2.66	554
	Between		0.18	0.94	1.57	
	Within		0.36	-1.14	2.62	
Economy	Overall	-0.49	0.73	-1.99	1.49	479
	Between		0.64	-1.48	0.87	
	Within		0.34	-1.62	0.46	

Note: Demand stands for energy demand (consumption), Oilprice represents crude oil price, urban is the degree of urbanisation whereas economy is the economic structure variable, ln denotes the natural log transformation of the variables

6.2.2 Correlation matrix and visual inspection of the variables

The term 'multicollinearity' is used to measure the degree of correlation among the independent variables in a model. Hence, the linear association between the variables used in the aggregate demand model is investigated by analysing the correlation among the variables. According to Gujarati (1988, p. 298, as citied in Tekaya, 2008), regarding multicollinearity

"it is a question of degree and not of kind. The meaningful distinction is therefore not the presence or the absence of multicollinearity, but between its various degrees".

Hence, the degree of strength of the relationship among the independent variables is used to determine the degree of multicollinerity in the model analysed. The correlation coefficients for the pairs of each variable series used in the aggregate demand model are provided in Table 6.2. From Table 6.2, it is evident from the results of the correlation coefficients that there is no problematic issue of multicollinearity among the variables analysed. In other words, since none of the coefficient value is higher than 0.75 using Tabachnich and Fidell's (2007) cut-off line, we do not have a multicollinearity problem in the model. The highest correlation found was between log of income and log of energy demand, with a correlation value of 0.60.

Table 6.2: Correlation matrix of the variables

	lnDemand	lnEconomy	lnUrban	lnIncome	lnOilprice
lnDemand	1.00				
lnEconomy	0.13	1.00			
lnUrban	-0.20	0.07	1.00		
lnIncome	0.60	0.37	-0.26	1.00	
lnOilprice	-0	-0.05	-0.20	0.07	1.00

After establishing that there is no multicollinearity problem in the model, we proceed to present the graphical plots (Figure 6.1 - 6.5) of each series of the variables used to analyse the driving forces of aggregate energy demand in Sub-Saharan Africa. This allows us to inspect the series visually, a practice that also allows us to detect any structural breaks in the series plotted. All the series for log of oil price for all countries appear to grow over time suggesting the presence of a unit root. However, for the series pertaining to the log of economic structure, urbanisation, income and energy demand, the evolution of the series suggests mean reversion and hence stationarity. It is also evident from the visual inspection of the plots that there are no obvious structural breaks. The numbers 1,2,3,…,16 in the plotted graphs represent series of the variable for each of the sixteen countries analysed in the panel model.

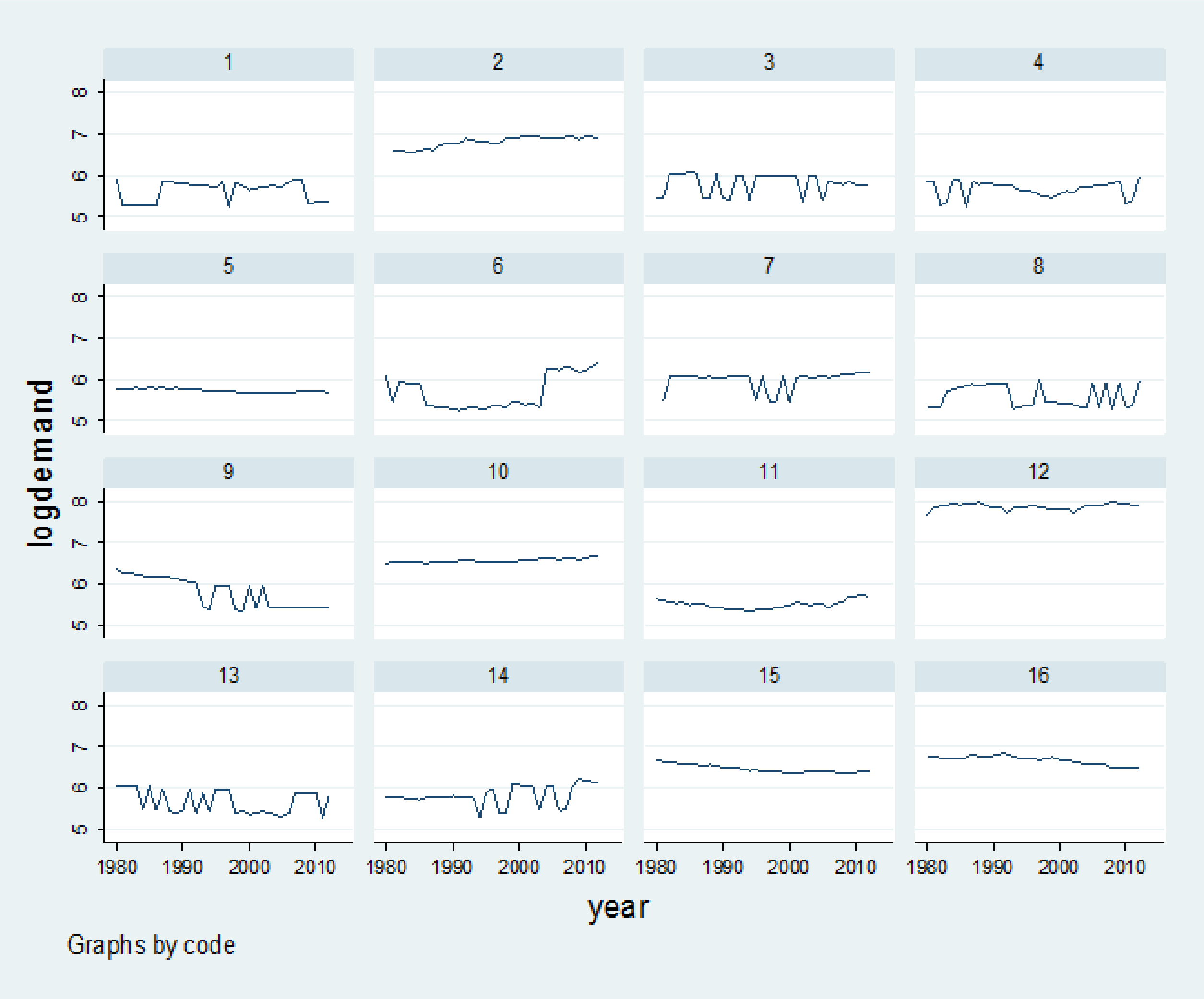

Figure 6.1: Graphical plots showing log of energy demand series

1,2,3,4,5,6….,16 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Mozambique, Nigeria, Senegal, South Africa, Sudan, Togo, Zambia and Zimbabwe log of aggregate energy consumption data series, respectively.

Figure 6.2: Graphical plots showing log of economic structure series

1,2,3,4,5,6….,16 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Mozambique, Nigeria, Senegal, South Africa, Sudan, Togo, Zambia and Zimbabwe log of economic structure data series, respectively.

Figure 6.3: Graphical plots showing log of urbanisation series

1,2,3,4,5,6….,16 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Mozambique, Nigeria, Senegal, South Africa, Sudan, Togo, Zambia and Zimbabwe natural log of degree of urbanisation data series, respectively.

Figure 6.4: Graphical plots showing log of income series

1,2,3,4,5,6….,16 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Mozambique, Nigeria, Senegal, South Africa, Sudan, Togo, Zambia and Zimbabwe natural log transformation of income data series, respectively.

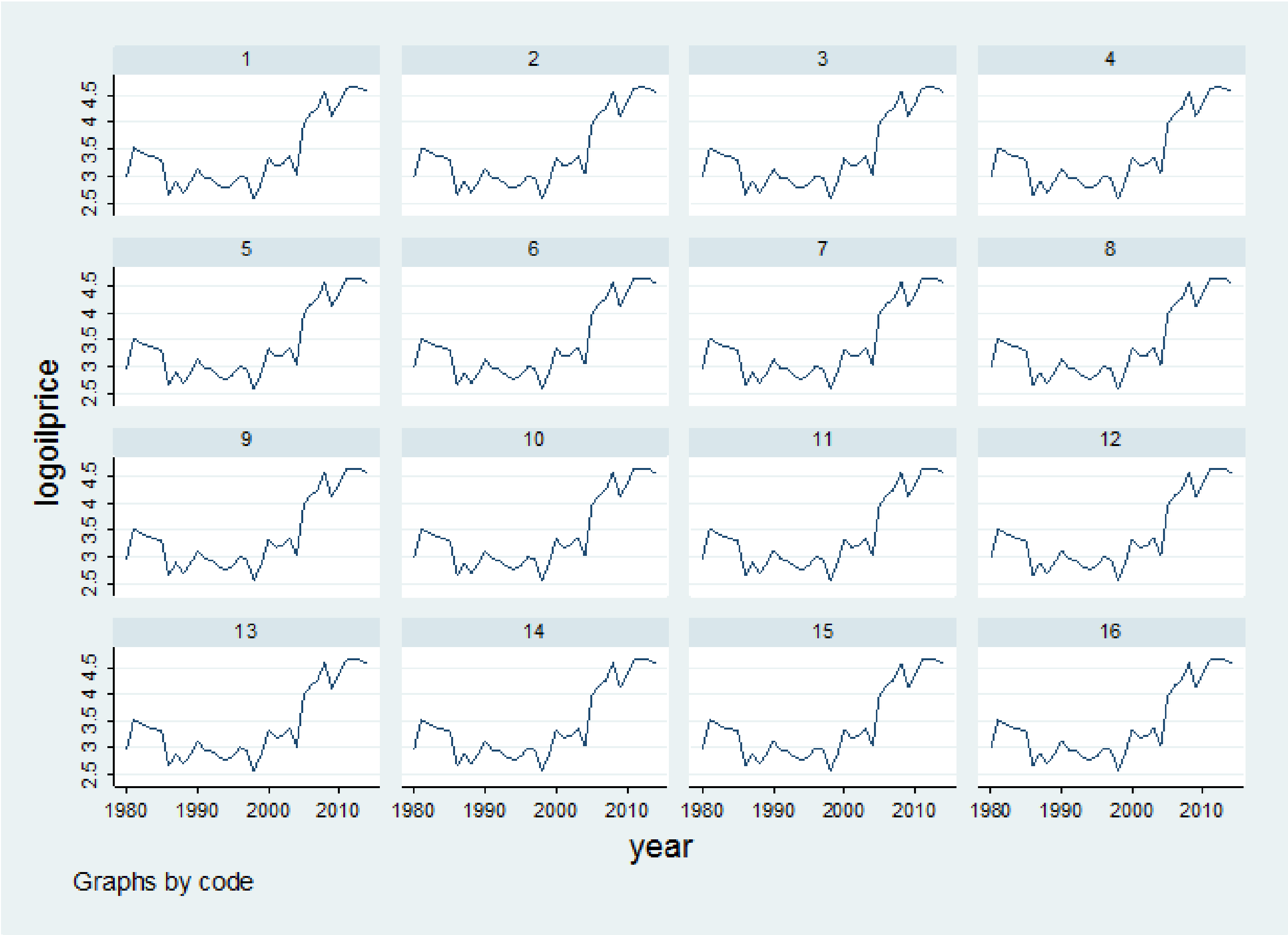

Figure 6.5: Graphical plots showing log of oil prices series

1,2,3,4,5,6….,16 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Mozambique, Nigeria, Senegal, South Africa, Sudan, Togo, Zambia and Zimbabwe natural log transformation of oil price data series, respectively.

The scatter plot of the relationship between energy demand and income is presented in Figure 6.6. The chart shows a clear positive relationship between the plotted variables. This suggests the existence of a positive long run relationship between the variables in question, a pattern that should find confirmation in our cointegration results to be presented below.

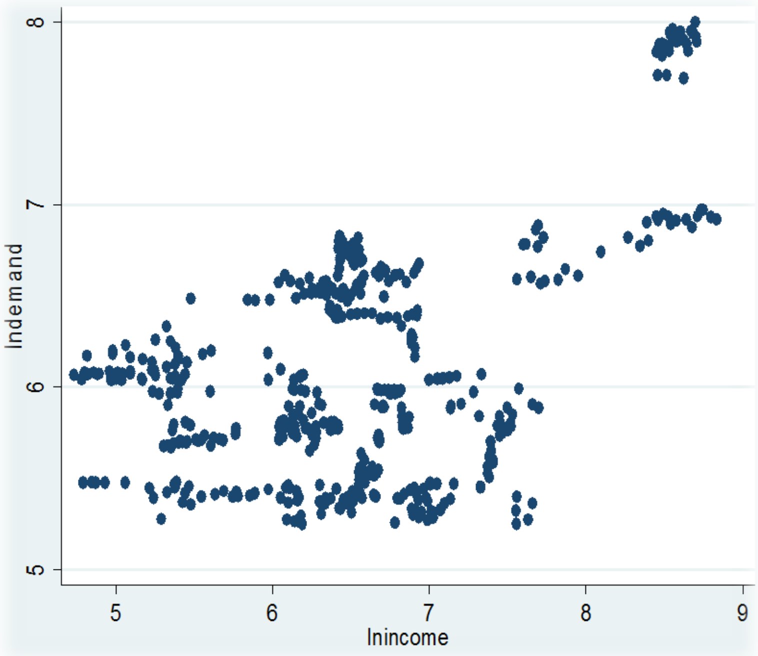

Figure 6.6: Scatter plot of energy demand and income

6.3 Panel unit root tests

Three panel unit root tests are employed to determine the stationarity properties of the variables in the model. That is, the Im-Pesaran-Shin unit root test and the two Fisher type (based on Augmented Dickey Fuller and the Phillips-Perron) panel unit root tests. The results of the panel unit root tests carried out to know the unit root properties of the variables are shown in Table 6.3. The tests chosen are those that give unbiased estimates even in an unbalanced panel (see Chapter 4).

Table 6.3: Panel unit root tests results for the variables in levels and first differences

Variable	IPS Statistics	ADF- Fisher	PP- Fisher	Inference
lnD	-3.069**	7.731***	13.391***	Stationary
lnY	2.979	-1.347	-0.867	Non-stationary
Eco	-2.771**	4.516***	6.479***	Stationary
Urb	-2.443*	6.085***	-1.046	Mixed
lnP	5.557	-3.775	-3.274	Non-stationary
ΔlnD	-16.597***	49.290***	86.474***	Stationary
ΔlnY	-11.454***	27.082***	64.513***	Stationary
ΔEco	-12.712***	34.301***	71.345***	Stationary
ΔUrb	-8.667***	16.540***	31.852***	Stationary
ΔlnP	-13.910***	0.999***	83.806***	Stationary

Notes: Δ is the first difference operator. *, ** and *** indicates rejection of the null of a unit root at the significance levels of 10%, 5% and 1%, respectively. lnD represents the natural log of energy demand, lnY stands for the natural log of income, Eco is the economic structure while Urb stands for the degree of urbanisation. Lastly, lnP is the natural log of crude oil price. The natural log of the degree of urbanisation and economic structure data series are not taken because they are in percentages.

As stated in the notes above, all three tests have the same null hypothesis of the series containing a unit root. The energy demand and economic structure variable series based on the three tests performed gave strong evidence that the variables are stationary both in levels and, obviously, in first difference. The log of income and oil price are non-stationary in levels, that is, they contain a unit root in levels but after their first difference is taken, the series become stationary. Lastly, the urbanisation variable gave evidence of stationarity for the series under the IPS and ADF- Fisher tests in levels but the results are in contrast with that of PP- Fisher panel unit root test which showed that the urbanisation series (variable) is non-stationary in levels. However, the mixed result was clarified when the first difference of the series was taken and all three tests gave strong evidence of stationarity. It is apparent from Table 6.3 that there is a mixed order of integration among the variables in levels.

Having established that there is a mixed order of integration among the variables, we can now proceed to investigate the long-run relationship between the variables, and whether there is in fact a statistically significant cointegrating relationship using the Westerlund-based panel cointegration test. Nonetheless, the works of Pesaran and Pesaran (1997), Pesaran *et al.* (2001) have shown that there exists a possibility of a long run cointegration relation among series of differing order of integration (that is I(0) and I(1)).

Base on this and following the approach of Frimpong and Adu (2014), Martins (2006), and Morshed (2010), we proceed to examine the cointegration relationship among the variables in this study. For instance, Frimpong and Adu (2014) examined the long run cointegration relationship between real GDP per capita, human capital, physical capital formation, openness and a set of human health indicators on a panel data set of 30 SSA countries between 1970-2010, using the Westerlund cointegration approach, and the PMG error correction technique to estimate the long and short run determinants of growth, in the presence of a mixed order of integration.

An advantage of using the PMG error correction approach especially in the context of this study is that the PMG technique is a panel extension of the Autoregressive Distributed Lag (ARDL) model popularly used in time series analysis due to its ability to estimate a cointegration relationship among variables even when they have different order of integration. The evidence above gives enough bases for this study, to proceed with the long run analysis even when the variables have different order of integration. The test was also chosen because it is robust enough to allow for high level of heterogeneity in the short run as well as in the long run cointegrating relationship, even in an unbalanced panel.

6.4 Panel cointegration test result

The Westerlund-based panel cointegration test is used to know whether the variables move together in the long run. In other words, the cointegration test is used to know whether energy demand, income, degree of urbanisation, oil price and economic structure are cointegrated. We specify a single lead and lag based on a constant and a trend using 150 replications through a bootstrapping procedure. Bootstrapping of the test statistics is used to correct for correlation among the cross-sectional units in the panel, thus giving robust critical values. The results obtained from the tests are presented in Table 6.4 below.

Table 6.4: Results of the Westerlund-based panel cointegration tests

Statistic	Value	Z-value	Robust P-value
Gt	-2.585	1.230	0.260
Ga	-8.611	4.050	0.070
Pt	-11.141	-0.632	0.010
Pa	-10.981	1.351	0.000

Notes: Gt and Ga represent the cross sectional statistics while Pt and Pa stands for the panel statistics. Since the bootstrap option is used, we use the robust *p*-value to make our decision. The null of the test is no cointegration among all the variables. The rejection of the null hypotheses (H0) would mean that there is cointegration in the panel.

As Table 6.4 shows, a robust *p*-value, under one of the cross sectional statistics (Ga) and the two panel statistics (Pt and Pa) give evidence that the null hypothesis of 'no cointegration' can be rejected. Rejection of one of the cross sectional statistics (Gt or Ga) is sufficient evidence against the null. Interestingly, both Pt and Pa robust *p*-values give strong evidence against the null hypothesis at the 1% significance level. Therefore, we have strong evidence of cointegration among the variables in our aggregate demand model. Since we have ascertained that there is cointegration among the variables, we can now safely proceed to estimate the long-run relationships among the variables.

6.5 Estimation and interpretation of the long run aggregate energy demand relationships

Having established cointegration, i.e., that the variables in the model do move together in the long-run forming an economically meaningful and statistically significant relationship, the analysis now proceeds to assess the long-run behaviour of the model in question. This section presents the results of the long run determinants of energy demand in Sub-Saharan Africa, based on the results of the Pooled Mean Group (PMG) estimator error correction model.

Table 6.5: Results of the long run estimates

Regressor	Elasticity result
lnIncome	0.099^{***}
	(0.000)
lnOilprice	-0.456^{**}
	(0.022)
Urban	0.013^{***}
	(0.007)
Ecostruc	-0.001^{*}
	(0.095)

Notes: The dependent variable is the log of energy demand, p values are in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

From Table 6.5, it is apparent that all the variables have the *apriori* expected sign except for the economic structure variable. The changes in price and income have the expected theoretical sign as discussed in section 6.2 above. Specifically, a 1% increase in consumers' income will lead to a 0.10% increase in energy consumption in SSA. The coefficient is strongly significant, at the 1% level, which implies that in the long run consumers will adjust and increase the amount of goods in their consumption bundles. This includes both energy and non-energy goods but it should be noted that the elasticity is relatively inelastic in the long run. Also, in the long run the appliance stock adjusted over time due to an increase in income level would be reflected by an increase in aggregate energy demand.

It is evident from this that the demand for energy is inelastic because the changes in income lead to a smaller increase in the amount of energy consumed. In fact, the change in the consumption level is very small when compared to the increase in the consumers income from the result reported.

The estimated coefficient of the energy price variable also gives evidence of a negative elasticity and it is significant at the 5% level. The estimated coefficient indicates that a 1% increase in energy price leads, on average, to a 0.46% reduction in the amount of energy consumed; a plausible result which is economically meaningful. It also confirms that energy demand is inelastic in the long-run in SSA as stated earlier, perhaps because energy is an essential good to most consumers.

Also, the estimated coefficient of the degree of urbanisation is statistically significant and the coefficient sign is as expected *a priori*. Energy demand appears to increase by 0.01% for every 1% increase in the population of urban areas in SSA. The small elasticity figure is likely to be the result of unavailability of key energy dependent infrastructure, like railways, in most SSA cities, which increase in urban population would have put pressure on. However, the small increase found may be due to the fact that when consumers move to urban areas, they move towards the use of more modern energy types, appliances and gadgets which may be linked to the increase in income level.

The estimated elasticity of the economic structure variable does not have the expected sign but it is statistically significant at the 10% level. A negative elasticity is found, which is likely to be due to the shift from industrial sector to service sectors in SSA in the last few decades. However, the result should be interpreted with caution as the elasticity found is very small (0.001%). The discussion of the results is presented in Section 7.2.

6.6 Estimation and interpretation of the short run aggregate energy demand relationships

The results of the estimated short run determinants of energy demand in SSA are presented and discussed in this section.

Table 6.6: Results of the ECM estimates in the short run

Regressor	Elasticity result
lnIncome	-0.059
	(0.682)
lnOilprice	0.238^{***}
	(0.00)
Urban	0.137^{**}
	(0.055)
Ecostruc	0.031
	(0.795)

Notes: The error correction term which measures the speed of convergence to long run equilibrium from the model analysis is -0.538^{***}, the dependent variable is the log of energy demand, *p*-values are in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

An important component of the error correction model (ECM) is the error correction term (henceforth, '*ec*'), which is expected to be negative and statistically significant. The result of this study confirms this because the *ec* is negative (-0.538) and statistically significant at the 1% level (*p*-value of 0.000). The *ec* shows that every year, 53.8% deviation of the variables from long run equilibrium is corrected, hence it takes just short of two years for short-run deviations or shocks to the variables to return to their 'equilibrium', i.e., long-run values.

All the variables except the income and price variables have the signs expected *a priori*. The results confirmed that for every 1% rise in urban population, there will be a significant increase of 0.14% in energy demand in SSA. Likewise, for every 1% increase in industrial output there will be an increase of 0.03% in energy consumption.

However, the result of the income elasticity is in contrast of the *a priori* expectation. This might be due to the low income levels in most of the analysed countries, and going by short run economic analysis, even when there is an increase in income it will take some time for consumers to adjust their consumption bundle to reflect this change in the short run. The same goes for the price variable which shows a positive price elasticity. This might also be because the demand for energy is inelastic and there are no close substitutes for most liquid fuels like kerosene and LPG, or consumers are less responsive to changes in price in the short run. But, considering that the income coefficient is not statistically significant, the result should be interpreted with caution. Besides, we are more concerned with the long run analysis which is the aim of most economic modelling like the present study.

Both the short-run and long-run results presented are supported by several studies in the literature, most of which were reviewed and discussed critically in Chapter 3. A thorough comparison of the results obtained in the present study with those available from prior literature will be undertaken in the discussion chapter that follows.

So far in the sections above, we have focused on the discussion of the estimated results of the determinants of aggregate energy demand in SSA. In the sections and subsections that follow the results of the analysed individual fuel types are presented and discussed.

6.7 Disaggregate energy demand analysis in Sub-Saharan Africa

The sections and sub-sections that follow reports the results of the disaggregated econometric analysis by energy types, by considering the demand for kerosene, petrol (gasoline), liquefied petroleum gas (LPG), biomass, diesel and electricity demand in SSA.

6.7.1 A priori expectations accounting for idiosyncracies of individual fuel types

Based on economic theory, the *a priori* expectations accounting for idiosyncracies of individual fuel types, are as follows:

H1: Positive income elasticity is expected because as the income level of consumers goes up, they increase the amount of commodities in their consumption bundle which includes the various energy types. This assumption is rooted in the demand theory discussed in Chapter Two. However, a negative income elasticity of solid biomass demand is expected. This is because people are expected to use modern energy as their income increases.

H2: Growth in urban population (urbanisation) is expected to have a positive impact on the aggregate demand for energy. This is because as people move from rural to urban areas, it is assumed that they change from the traditional forms of energy like solid biomass to modern energy types like electricity and LPG.

H3: A positive relationship is expected between energy demand and economic structure. Considering that a country economic structure is derived by using the share of industrial value added to the service value added, which is used to measure the impact of structural changes in the countries. An increase in industrial share or output will increase the amount of energy consumed. However, for the biomass demand model, a negative relationship is expected because more industrial output can lead to the production of affordable modern energy cooking stoves.

H4: An increase in population will increase the amount of energy used in SSA by the consumers for all fuel types. Therefore, a positive relationship between energy demand and population is expected.

6.7.2 Descriptive statistics of the variables used in the disaggregated analysis

The table below (Table 6.7) presents the basic characteristics of the data used to analyse the main determinants of different energy types in SSA. In other words, as presented earlier for the aggregate analysis, the table below gives a brief but informative summary of the dataset used. For instance, from the table we see that the total number of observations under urban area growth is 408 from the 12 countries analysed in the 33 years period.

Table 6.7: Summary statistics of the disaggregate energy demand in SSA dataset

Variable		Mean	Std Dev	Min	Max	Obs
lnIncome	Overall	6.65	1.00	4.74	8.92	407
	Between		0.12	6.45	6.95	
	Within		0.10	4.85	8.66	
lnUrban	Overall	1.34	0.40	0.09	2.66	408
	Between		0.17	1.15	1.62	
	Within		0.36	0.07	2.42	
lnEconomy	Overall	-0.44	0.76	-1.70	1.49	406
	Between		0.10	-0.67	-0.26	
	Within		0.75	-1.96	1.45	
lnPop	Overall	16.42	1.28	13.81	18.97	408
	Between		0.27	15.96	16.85	
	Within		1.25	14.09	18.58	
lnKero	Overall	3.80	1.73	0.00	7.82	404
	Between		0.20	3.46	4.16	
	Within		1.72	-0.13	7.90	
lnBiomass	Overall	11.67	1.62	8.93	15.23	406
	Between		0.22	11.31	12.07	
	Within		1.61	8.92	14.91	
lnDiesel	Overall	5.85	1.30	1.95	9.15	406
	Between		0.36	5.38	6.56	
	Within		1.25	2.42	8.83	
lnLPG	Overall	2.78	1.64	0.00	6.14	351
	Between		0.46	2.14	3.57	
	Within		1.58	-0.59	6.16	
lnPetrol	Overall	5.49	1.56	2.30	9.05	406
	Between		0.37	5.04	6.16	
	Within		1.52	2.61	9.31	
lnElect	Overall	7.71	1.73	4.23	12.25	406
	Between		0.44	7.02	8.57	
	Within		1.68	4.84	12.18	

Note: ln stands for the natural log transformation of the variables, lnUrban is the natural log transformation of urbanisation, lnEconomy is the natural log transformation of economic structure, lnPop is the natural log transformation of Population, lnKero is the natural log transformation of kerosene consumption, lnBiomass is the natural log transformation of biomass consumption, lnDiesel is the natural log transformation of diesel consumption, lnLPG is the natural log transformation of LPG consumption, lnPetrol is the natural log transformation of petrol consumption and lnElect is the natural log transformation of electricity consumption. The number of observations under LPG reduced from 406 to 351 after taking the log transformation because some of the values in the variables series of some countries had value of 0's.

As illustrated in Table 6.7, it is evident that we have an almost balanced panel which informed the selection of the model used for the analysis. That is, only four values are missing for most series in the dataset used. This can be accommodated by the software (STATA 13) used for the analysis, an adjustment that ensures that the estimates are not biased.

Having presented the summary statistics of the dataset used, what follows is the visual inspection of the plots of each variable to detect any structural break in the series. The number on each plot in the charts represents each series of the plotted variables for the countries analysed.

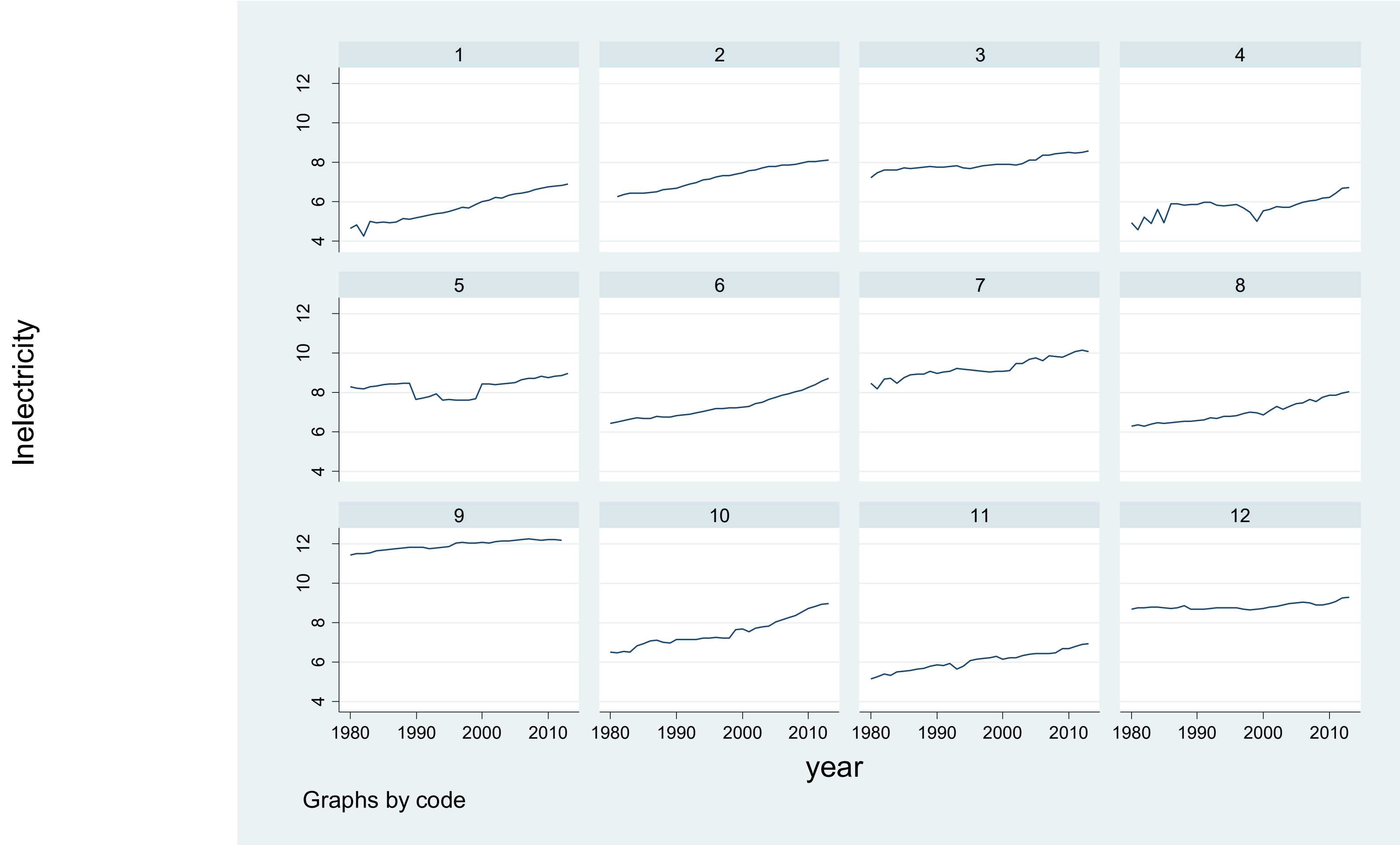

Figure 6.7: Graphical plots showing log of electricity demand series

1,2,3,4,5,6….,12 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Nigeria, Senegal, South Africa, Sudan, Togo and Zambia log of electricity consumption data series, respectively.

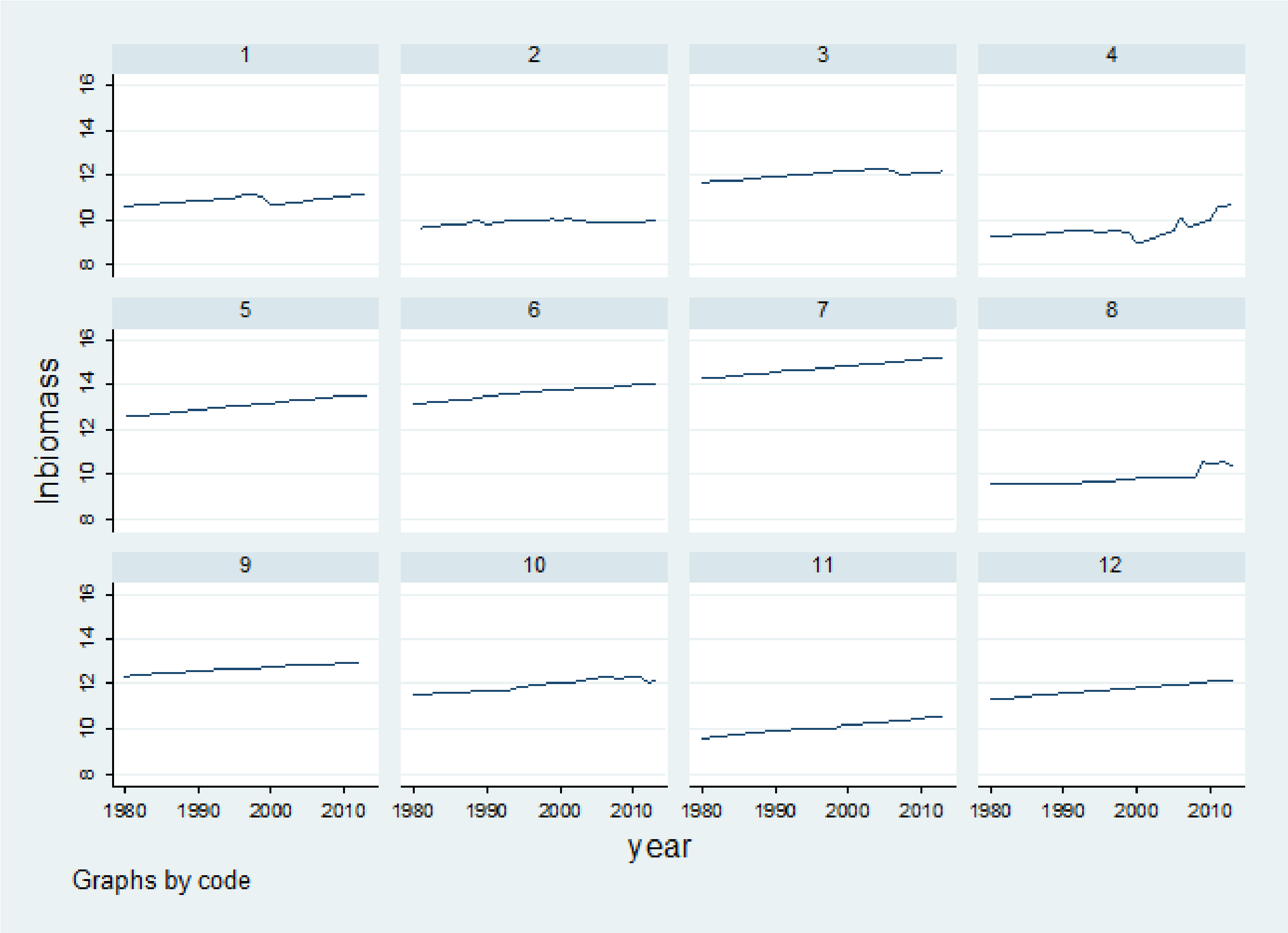

Figure 6.8: Graphical plots showing log of biomass demand series

1,2,3,4,5,6….,12 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Nigeria, Senegal, South Africa, Sudan, Togo and Zambia biomass consumption data series, respectively.

Figure 6.9: Graphical plots showing log of petrol demand series

1,2,3,4,5,6….,12 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Nigeria, Senegal, South Africa, Sudan, Togo and Zambia petrol consumption data series, respectively.

Figure 6.10: Graphical plots showing log of kerosene demand series

1,2,3,4,5,6….,12 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Nigeria, Senegal, South Africa, Sudan, Togo and Zambia kerosene consumption data series, respectively.

Figure 6.11: Graphical plots showing log of diesel demand series

1,2,3,4,5,6….,12 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Nigeria, Senegal, South Africa, Sudan, Togo and Zambia diesel consumption data series, respectively.

Figure 6.12: Graphical plots showing log of LPG demand series

1,2,3,4,5,6….,12 on the plots represent: Benin, Botswana, Cameroon, Congo, Democratic Republic of Congo, Ethiopia, Nigeria, Senegal, South Africa, Sudan, Togo and Zambia LPG consumption data series, respectively.

From the visual inspection of Figures 6.7 to 6.12, we cannot detect any pronounced structural breaks in the series analysed. In other words, the visual inspection of the plots suggests no major or persistent structural breaks. We, therefore, proceed to present the results of each of the analysed energy types.

6.8 Panel unit root tests

Three panel unit root tests are employed to determine the stationarity properties of the variables in the model. That is, the Im-Pesaran-Shin unit root test and the two Fisher type (based on Augmented Dickey Fuller and the Phillips-Perron) panel unit root tests. The results of the panel unit root tests carried out to know the unit root properties of the variables employed in the disaggregated analysis by fuel type are shown in Table 6.8 below.

Table 6.8: Panel unit root tests results for the variables in levels and first differences

Variable	IPS Statistics	ADF- Fisher	PP- Fisher	Inference
Lnkero	0.187	-0.0310	-0.031	Non-stationary
lnPet	1.497	1.117	1.117	Non-stationary
lnLPG	-0.418	-1.342	-1.342	Non-stationary
lnBio	2.073	1.651	1.651	Non-stationary
lnElec	5.712	3.401	3.401	Non-stationary
lnDie	3.851	3.832	3.832	Non-stationary
ΔlnKero	-9.564^{***}	-17.128^{***}	-17.128^{***}	Stationary
ΔlnPet	-8.899^{***}	-15.909^{***}	-15.357^{***}	Stationary
ΔlnLPG	-9.176^{***}	-15.909^{***}	-15.909^{***}	Stationary
ΔBio	-8.453^{***}	-15.385^{***}	-15.385^{***}	Stationary
ΔlnElec	-13.910^{***}	-18.881^{***}	-18.881^{***}	Stationary
ΔlnDie	-9.950^{***}	-17.183	-17.183^{***}	Stationary

Notes: Δ is the first difference operator. *, ** and *** indicate rejection of the null of a unit root at the significance levels of 10%, 5% and 1%, respectively. lnkero represents the natural log of kerosene demand, lnPet stands for the natural log of petrol demand. lnLPG is the log of LPG demand. lnBio is the log of solid biomass demand, lnElec is the log of electricity demand, lnDie is the log of diesel demand.

As stated in the table notes above, all three tests have the same null hypothesis of a unit root. All variables are non-stationary or contain a unit root in levels whereas at first difference they are all stationary (i.e., integrated of order zero, or I(0)). In other words, when the unit root properties of the series were taken in levels, no evidence was found against the null hypothesis. This led to the conclusion that the series contain unit roots.

However, after taking the first difference of the variables and the unit root tests were performed, strong evidence was found at the 1% significance level against null hypothesis. We can, therefore, conclude that the series do not contain unit root after first differencing. Further, since all variables are stationary at first difference this would help reduce any serial correlation that may be present in the series.

6.9 Electricity demand model

The results of the linear panel models used to analyse the determinants of electricity demand in SSA are presented in this part of the thesis. The correlation matrix of the model is also presented to rule out any serious issue of multicollinearity. Also, the chart of two selected variables from the model is also highlighted.

Table 6.9: Correlation matrix of the variables

	lnIncome	lnUrban	lnEconomy	lnPopulation	lnElectricity
lnIncome	1.00				
lnUrban	-0.30	1.00			
lnEconomy	0.33	0.11	1.00		
lnPopulation	-0.32	0.02	-0.01	1.00	
lnElectricity	0.40	-0.33	0.09	0.62	1.00

The results, as shown in Table 6.9, indicate that there is no problem of multicollinearity among the variables analysed. Since none of the coefficient value is higher than 0.75 using the Tabachnich and Fidell's (2007) cut-off line, we do not have a multicollinearity issue. The highest correlation was between log of electricity and log of population (equal to 0.62).

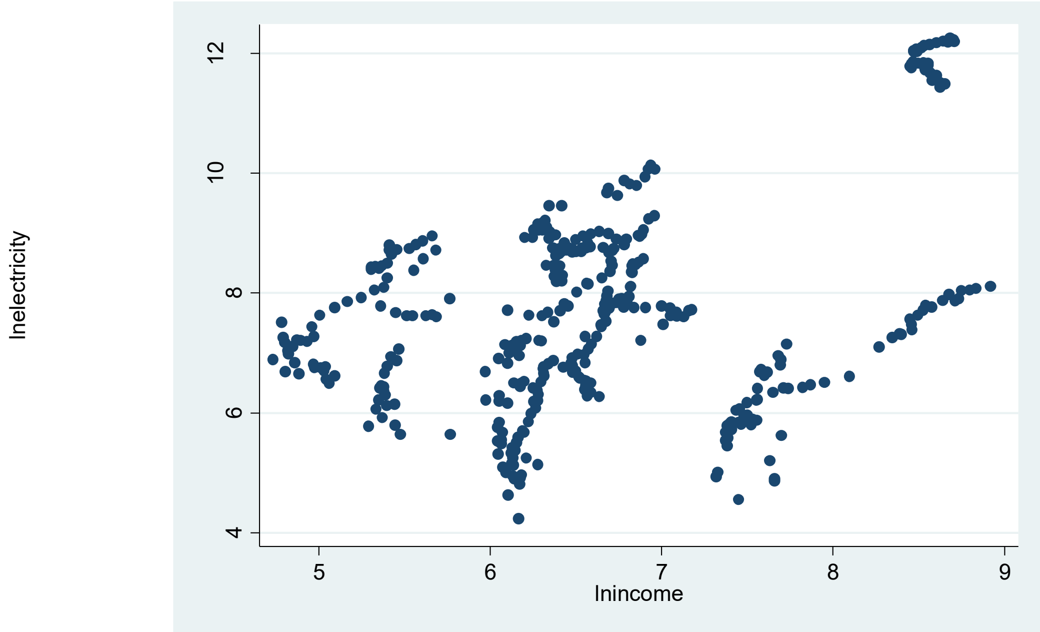

Figure 6.13: Scatter plot of electricity demand and income

As illustrated in Figure 6.13, there is a positive relationship between the two variables plotted. The plot suggests that a positive relationship exists between electricity and income. The observed pattern from the chart should be corroborated and explained by our regression results below.

Table 6.10: Estimation results for electricity demand

Dependent variable: log of electricity	Panel A: Fixed Effects		Panel B: Random Effects		Panel C: PW Model	
Explanatory variables	Coeff	p-value	Coeff	p-value	Coeff	p-value
lnIncome	0.545***	0.000	0.585***	0.000	0.595***	0.000
inUrbanisation	-0.113***	0.002	-0.121***	0.009	-0.242***	0.009
lnEconomic structure	0.089***	0.010	0.085**	0.016	-0.009	0.811
lnPopulation	1.537***	0.000	1.487***	0.000	1.015***	0.000
Constant	-20.951***	0.000	-20.39***	0.000	-12.583***	0.000
F test for model significant						
F-statistic for fixed effects and Wald statistic for random effects and PW	447.24		1714.35		343.60	
P-value>F	0.000		0.000		0.000	
Observations	404		404		404	
Groups	12		12		12	
Model goodness-of-fit						
R-squared					0.8239	
Within	0.822		0.821			
Between	0.595		0.618			
Overall	0.610		0.631			
Autocorrelation test (Wooldridge test)			1.492			
Heteroscedasticity test (Wald test)	869.52***					
Hausman Test		Fixed vs Random effects				
Test statistic	54.02					
p-value> Test statistic	0.000					
Decision		Fixed effects model				

Note: *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$, PW represents Prais-Winsten estimation with heteroscedastic panel-corrected standard errors, ln stands for natural logarithm.

The fixed effects model is appropriate for the analysis of electricity demand according to the Hausman selection test carried out, as shown in Table 6.10 above. In the Hausman test, the chi-squared critical value is lower than the test statistic for the individual effect model, therefore, the fixed effects model is preferable. This is, going with the test statistic of using fixed effects model if the p-value is < 0.05, and in the analysis the result we had recorded a p-value of

0.000. Since, the p-value of 0.000 is less than 0.05, we can safely accept the hypothesis that the fixed effects model is appropriate.

However, despite the selection of the fixed effects model, its regression estimates may not be the best linear unbiased estimator (BLUE) because of the evidence of the presence of heteroscedasticity. The modified Wald test for groupwise heteroscedasticity in fixed effects model carried out has a null of constant variance (homoscedasticity). We had a *p*-value of 0.000, which led to the rejection of the null hypothesis and conclusion of heteroscedasticity in the model.

Interestingly, from the result of the Wooldridge's test for autocorrelation, we accepted the null and rejected the alternative hypothesis of autocorrelation in the model. In other words, from the test results, we had a p-value of 0.248, which supported the null hypothesis of no serial correlation. This would suggest that the random effects model estimates may be efficient if chosen by the Hausman test. However, since the fixed effect model was chosen we need to use an estimator that can correct for heteroscedasticity found in the model. To correct for this, we use the PW model so as to obtain more efficient estimates.

According to the PW model (see the last two columns of Table 6.10 for the results), the income elasticity of electricity demand is positive, inelastic and strongly significant. That is, a 1% increase in income increases electricity consumption by approximately 0.60%. But the urbanisation elasticity is against our *a priori* assumptions. There was no increase in electricity consumption associated with urbanisation. Instead, we find that a unit increase in urbanisation reduces electricity consumption by 0.24%, and the estimated coefficient is significant at the 1% statistical level.

The results also show that population is a key and significant factor behind electricity consumption in SSA. Specifically, a 1% increase in population size is expected to lead to a 1.02% increase in energy demand. On the other hand, economic structure is not found to be a statistically significant factor driving electricity consumption in SSA.

These results suggest that the demand for electricity is relatively inelastic to changes in income and urbanisation, while it is has unit elasticity to changes in the size of the population.

6.10 Diesel demand model

The results of the correlation matrix of the variables and scatter plot of two key variables are first presented below, before moving on to present the table showing the estimates from the fixed effects model, random effects model and the PW model employed to investigate the driving forces of diesel demand in SSA.

Table 6.11: Correlation matrix of the variables

	lnIncome	lnUrban	lnEconomy	lnPopulation	lnDiesel
lnIncome	1.00				
lnUrban	-0.30	1.00			
lnEconomy	0.33	0.11	1.00		
lnPopulation	-0.32	0.02	-0.01	1.00	
lnDiesel	0.32	-0.25	0.09	0.70	1.00

As can be seen from the table above, there is no issue of multicollinearity among the variables analysed, according to the Tabachnich and Fidell's (2007) cut-off line of 0.75. The highest correlation was between log of diesel demand and log of population (with a recorded correlation coefficient of 0.70).

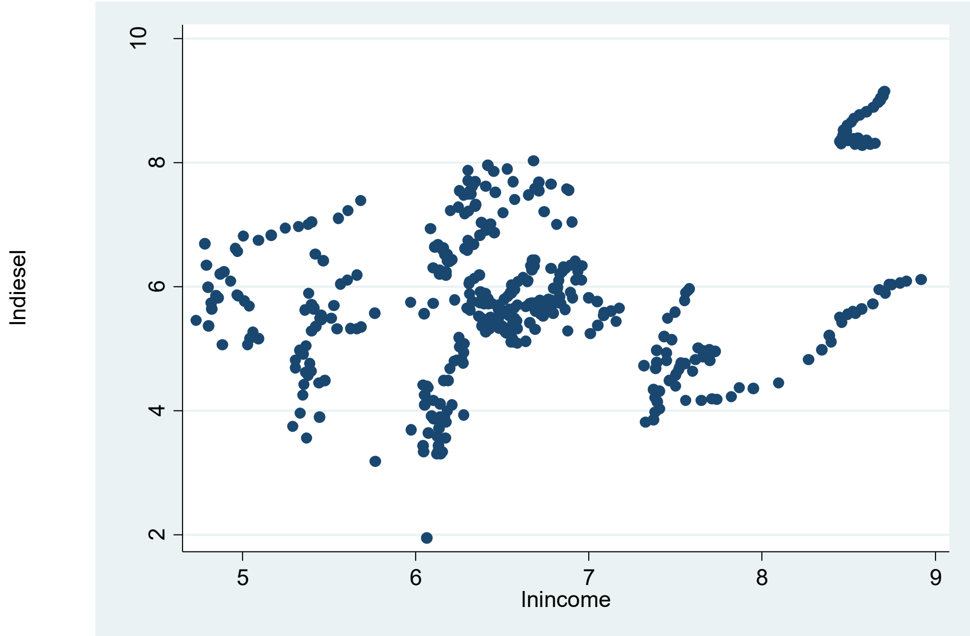

Figure 6.14: Scatter plot of diesel demand and income

As can be seen from Figure 6.14, the slope suggests an upward trend. Hence, the plot suggests that a positive relationship exists between diesel demand and income in SSA. Table 6.12 shows the results of the linear panel models used to analyse diesel demand in SSA. The results confirm the observed pattern illustrated by Figure 6.14.

Table 6.12: Estimation results for diesel demand models

Dependent variable: log of diesel	Panel A: Fixed Effects		Panel B: Random Effects		Panel C: PW Model	
Explanatory variables	Coeff	p-value	Coeff	p-value	Coeff	p-value
lnIncome	0.470***	0.000	0.567***	0.000	0.290***	0.001
lnUrbanisation	-0.214**	0.002	-0.235***	0.001	-0.350***	0.000
lnEconomic structure	0.029***	0.000	0.174**	0.004	0.043	0.384
lnPopulation	1.156***	0.000	1.002***	0.000	0.762***	0.000
Constant	-15.880***	0.000	-13.998***	0.000	-8.049***	0.000
F test for model significant						
F-statistic for fixed effects and Wald statistic for random effects and PW	82.54		402.70		137.81	
P-value>F	0.000		0.000		0.000	
Observations	404		404		404	
Groups	12		12		12	
Model goodness-of-fit						
R-squared					0.6687	
Within	0.4597		0.4532			
Between	0.7813		0.8447			
Overall	0.7171		0.7707			
Autocorrelation test (Wooldridge test)			41.805***			
Heteroscedasticity test (Wald test)	9475.62***					
Hausman Test	Fixed vs Random effects					
Test statistic	19.59					
p-value> Test statistic	0.000					
Decision	Fixed effects model					

Note: *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$, PW represents Prais-Winsten estimation with heteroscedastic panel-corrected standard errors, ln stands for natural logarithm

As shown in Table 6.12, the fixed effects model is the best model for analysing the determinants of diesel demand in SSA. This is supported by the result of the Hausman test with the chi-squared critical value being lower than the test statistics for the individual effect model; therefore the fixed effects model is preferable. Further, the p-value of 0.000 gives strong evidence in favour of the use of the fixed effects model.

The diesel fixed effects demand model suffers from heteroscedasticity and autocorrelation. To be specific, we strongly rejected the null of no serial correlation at the 1% significance level in the Wooldridge test for autocorrelation. While in the Wald test, the null of homoscedasticity was also strongly rejected due to a *p*-value of 0.000. To correct for these issues we used the PW model.

The PW model suggests that income, urbanisation and population are the key determinants of diesel demand in SSA. Their elasticities are 0.29%, -0.35% and 0.76%, respectively. To be specific, the income elasticity was found to be positive, inelastic and statistically significant at the 1% level. For every 1% increase in the size of urban population, diesel demand reduces by 0.35%. All the variables have the expected sign except for 'urbanisation' which has a negative sign.

6.11 Liquefied petroleum gas (LPG) demand model

The correlation matrix of the variables, scatter plot showing the relationship between the log of LPG and the log of income, results of the fixed effects model, random model and the Prais-Winsten model for LPG demand in SSA, are presented below.

Table 6.13: Correlation matrix of the variables

	lnIncome	lnUrban	lnEconomy	lnPopulation	lnLPG
lnIncome	1.00				
lnUrban	-0.30	1.00			
lnEconomy	0.31	0.19	1.00		
lnPopulation	-0.31	0.05	-0.09	1.00	
lnLPG	0.48	-0.25	0.02	0.44	1.00

From Table 6.13 above, we do not have a problem of multicollinearity among the variables analysed. According to the Tabachnich and Fidell's (2007) cut-off line of 0.75, we do not have a multicollinearity issue in the model because the highest correlation coefficient is 0.48, which is below 0.75. This is the coefficient of the relationship between the log of LPG demand and log of income.

What follows is the scatter plot (Figure 6.15) which shows that a positive relationship exists between income and LPG demand in SSA.

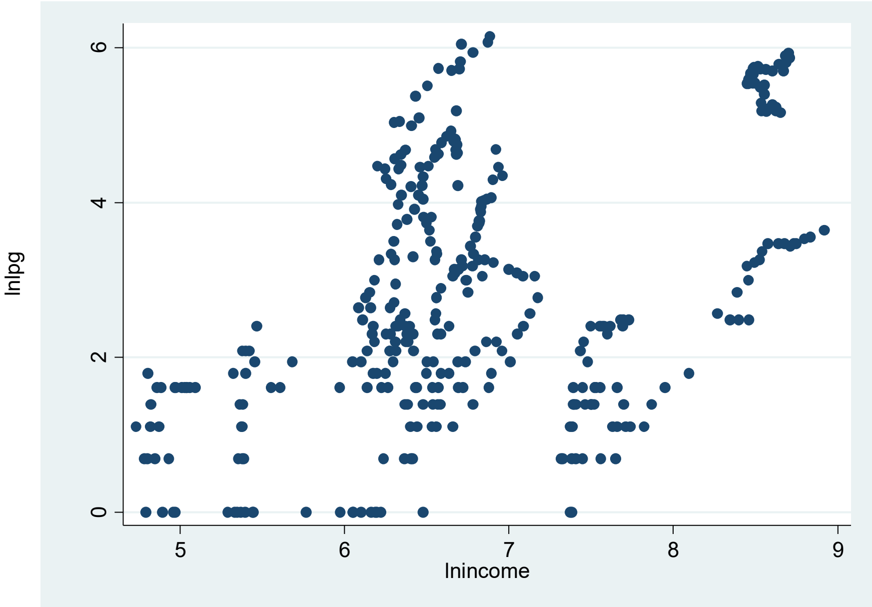

Figure 6.15: Scatter plot of LPG demand and income

Figure 6.15 suggests that when income increases, the demand for LPG also increases. This should be confirmed by the linear panel models used to analyse diesel demand in SSA presented and discussed below.

Table 6.14: Estimation results for LPG demand models

Dependent variable: log of LPG	Panel A: Fixed Effects		Panel B: Random Effects		Panel C: PW Model	
Explanatory variables	Coeff	p-value	Coeff	p-value	Coeff	p-value
lnIncome	0.884***	0.000	1.029***	0.000	1.097***	0.000
lnUrbanisation	-0.364***	0.000	-0.391***	0.000	-0.057	0.631
lnEconomic structure	0.755***	0.000	0.662***	0.000	-0.158	0.403
lnPopulation	1.623***	0.000	1.400***	0.000	0.948***	0.000
Constant	-28.942***	0.000	-26.332***	0.000	-20.264***	0.000
F test for model significant						
F-statistic for fixed effects and Wald statistic for random effects and PW	102.58		385.65		103.65	
P-value>F	0.000		0.000		0.000	
Observations	385		385		385	
Groups	12		12		12	
Model goodness-of-fit						
R-squared					0.1106	
Within	0.527		0.523			
Between	0.195		0.222			
Overall	0.241		0.274			
Autocorrelation test (Wooldridge test)			34.314***			
Heteroscedasticity test (Wald test)	4560.03***					
Breusch- Pagan selection test	1307.74***					
Breusch- Pagan test decision	Random effects model over OLS					
Hausman Test	Fixed vs Random effects					
Test statistic	15.18					
p-value> Test statistic	0.0043					
Decision	Fixed effects model					

Note: *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$, PW represents Prais-Winsten estimation with heteroscedastic panel-corrected standard errors, ln stands for natural logarithm, OLS represents Ordinary Least Square Model

From Table 6.14, we can see that the fixed effects model is chosen over the random effects model by the Hausman test. This is supported by the test statistic of 0.43% which is lower than 5% used in favour of using random effects model when deciding between using the random effects model or the fixed effects model. In other words, the fixed effects model is preferable because the chi-squared critical value is lower than the test statistics for the individual effect model.

The autocorrelation and heteroscedasticity tests carried out show that there is both autocorrelation and heteroscedasticity in the model. To be specific, we accepted the null hypothesis that there is autocorrelation in the model at the 1% significance level. Also, the null hypothesis of homoscedasticity was rejected due to strong evidence at the significance level of 1%. We therefore use the results of the PW model in the analysis and discussion that follows.

In the PW model, the main factors that drive LPG demand are income and population. Both have the expected positive sign, are moderately elastic and statistically significant. Urbanisation is also negative in this model like the one reported in the fixed effects model but it is not significant. Likewise, a negative relationship is found between LPG demand and economic structure which is against *a priori* expectations.

The high income elasticity found in the LPG model might be because, as consumers have more income, they tend to move toward modern energy like LPG, away from traditional sources like solid biomass. According to the model, a 1% increase in income induces an increase of 1.10% in LPG demand in SSA.

6.12 Petrol demand model

This section presents the results of the linear panel models used to analyse the petrol demand model in SSA. The correlation matrix of the model to rule out any serious issue of multicollinearity among the variables employed for the analysis is reported in Table 6.15.

Table 6.15: Correlation matrix of the variables

	lnIncome	lnUrban	lnEconomy	lnPopulation	lnPetrol
lnIncome	1.00				
lnUrban	-0.30	1.00			
lnEconomy	0.33	0.11	1.00		
lnPopulation	-0.32	0.02	-0.01	1.00	
lnPetrol	0.40	-0.18	0.18	0.64	1.00

From Table 6.15, it is evident that there is no issue of multicollinearity among the variables analysed. Since none of the coefficient values is higher than 0.75 using the Tabachnich and Fidell's (2007) cut-off line, we do not have a multicollinearity problem.

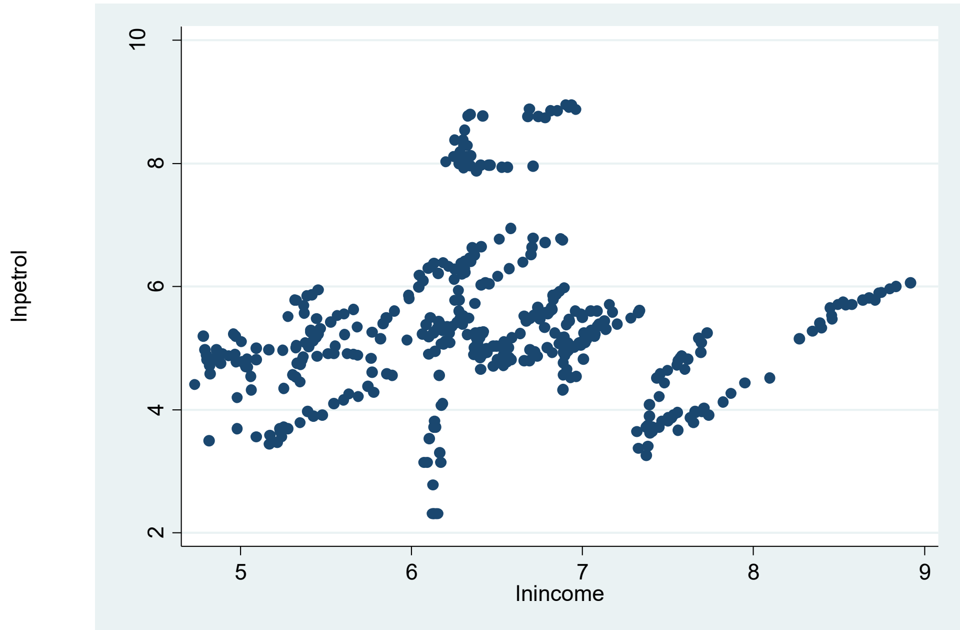

Figure 6.16: Scatter plot of petrol demand and income

As shown in Figure 6.16, there appears to be a positive relationship between the demand for petrol and consumers' income. This would suggest that as the income of consumers' increases, they demand more petrol as part of their consumption bundle.

Table 6.16: Estimation results for petrol demand models

Dependent variable: log of petrol	Panel A: Fixed Effects		Panel B: Random Effects	
Explanatory variables	Coeff	p-value	Coeff	p-value
lnIncome	0.548***	0.000	0.630***	0.000
lnUrbanisation	-0.229***	0.001	-0.227***	0.001
lnEconomic structure	0.164**	0.012	0.159***	0.012
lnPopulation	1.176***	0.000	1.115***	0.000
Constant	-17.088***	0.000	-16.638***	0.000
F test for model significant				
F-statistic for fixed effects and Wald statistic for random effects and PW	90.820		405.050	
P-value>F	0.000		0.000	
Observations	404		404	
Groups	12		12	
Model goodness-of-fit				
R-squared				
Within	0.484		0.482	
Between	0.729		0.774	
Overall	0.695		0.734	
Autocorrelation test (Wooldridge test)			31.087***	
Heteroscedasticity test (Wald test)	4081***			
Breusch- Pagan selection test	2134.18***			
Breusch- Pagan test decision	Random effects model over OLS			
Hausman Test	Fixed vs Random effects			
Test statistic	8.27			
p-value> Test statistic	0.082			
Decision	Random effects model			

Note: *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$

According to the Breusch-Pagan and the Hausman selection tests reported in Table 6.16, the random effects model is best for analysing the determinants of petrol demand in SSA. For the Breusch-Pagan test, we had strong evidence against the null hypothesis stating that pooled regression model is appropriate, hence we accept the alternative hypothesis that random effects model is appropriate. Likewise in the Hausman test, the chi-squared critical value is higher than the test statistics for the individual effect model; therefore the random effects model is preferable.

We do not need to analyse nor discuss the fixed effects model since it was not selected by the diagnostic tests. Moreover, the Wald test gave strong evidence of heteroscedasticity in the model which suggests that the estimates from the fixed effects model may be biased.

We had evidence of serial correlation in the random effects model going by the Wooldridge's test reported in Table 6.16. The null hypothesis of the test is that there is autocorrelation in the panel model. We had a strong test statistic from the analysis at the 1% significance level (p-value of 0.000), which implies that we had no reason to reject the null hypothesis. This does not pose any serious problem because since all variables are stationary at first difference, this would help reduce the serial correlation (Keneko, 2010). Therefore, we proceed to analyse and discuss the results of the random effects model.

As shown in the last two columns of Table 6.16, all the variables have the expected sign except for the urbanisation variable. Specifically, a 1% increase in consumers' income increases petrol demand by 0.63% in SSA, this is in line with the *a priori* expectation. This estimated coefficient is statistically significant at the 1% significance level. Also, a 1% increase in urbanisation reduces energy demand by 0.23%, and it is also statistically significant. A 1% increase in industrial output is expected to increase petrol demand by 0.16%. Lastly, energy demand is expected to increase by 1.12% for every percentage increase in population size in SSA. The constant has no economic interpretation as its result in this model only suggests that the constant term is different from zero (0).

Overall, these results indicate that the demand for petrol is relatively inelastic in SSA especially in response to changes in income, price and economic structure. However, petrol demand has a unit elasticity to changes in population size. All results are supported by several studies in the literature (see Chapter 3 and 7).

6.13 Kerosene demand model

As can be seen from Table 6.17 below, we do not have a problem of multicollinearity because the highest correlation coefficient from the results is 0.59. This is from the relationship between log of kerosene and log of population.

Table 6.17: Correlation matrix of the variables

	lnIncome	lnPopulation	lnUrban	lnEconomy	lnKerosene
lnIncome	1.00				
lnPopulation	-0.32	1.00			
lnUrban	-0.30	0.02	1.00		
lnEconomy	0.33	-0.01	0.11	1.00	
lnKerosene	0.13	0.59	0.02	0.13	1.00

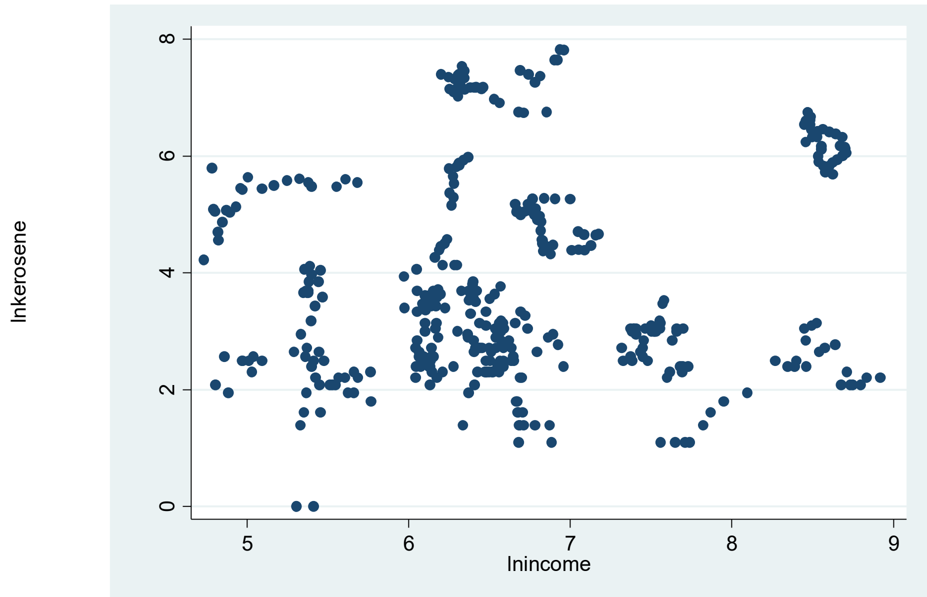

Figure 6.17: Scatter plot of kerosene demand and income

Figure 6.17 shows a positive relationship between the plotted variables, a pattern that should find confirmation in our regression results reported below.

Table 6.18: Estimation results for kerosene demand models

Dependent variable: log of kerosene	Panel A: Fixed Effects		Panel B: Random Effects	
Explanatory variables	Coeff	p-value	Coeff	p-value
lnIncome	0.735***	0.000	0.691***	0.000
lnUrbanisation	0.688***	0.001	0.585***	0.001
lnEconomic structure	-0.049	0.607	-0.021	0.823
lnPopulation	0.976***	0.000	0.960***	0.000
Constant	-18.066***	0.000	-17.356***	0.000
F test for model significant				
F-statistic for fixed effects and Wald statistic for random effects and PW	83.81		351.35	
P-value>F	0.000		0.000	
Observations	402		402	
Groups	12		12	
Model goodness-of-fit				
R-squared				
Within	0.479		0.479	
Between	0.167		0.209	
Overall	0.469		0.470	
Autocorrelation test (Wooldridge test)			6.872***	
Heteroscedasticity test (Wald test)	419.28***			
Breusch- Pagan selection test	2426.64***			
Breusch- Pagan test decision	Random effects model over OLS			
Hausman Test	Fixed vs Random effects			
Test statistic	4.14			
p-value> Test statistic	0.3873			
Decision	Random effects model			

Note: *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$

In deciding the appropriate model between fixed effects model and random effects model, the Hausman test gave evidence that the random effects model works well in analysing the determinants of kerosene demand in SSA.

The selection of random effects model might be influenced by the type of energy type that is analysed here. Considering that kerosene is used by low income earners and most of them are in the rural areas, it might suggest that it includes informal sectors which will be difficult to record by the national statistical offices during the data collection process. Therefore, the data reported might be a random selection from the population or estimation. This does not lead to any bias in our analysis, as it only confirms what is expected with an energy product like kerosene.

Therefore, we use the random effects model for our analysis because it is efficient for analysing the data even in the presence of autocorrelation. As discussed earlier since we have ascertained that the variables are stationary in first difference, any autocorrelation present in the model will not make the reported estimates less efficient.

According to the estimated model (see Table 6.18), income, urbanisation and population are the main factors that influence the demand for kerosene in SSA. To be specific, a 1% increase in income increases kerosene demand by 0.69% (significant at the 1% statistical level). Likewise, an increase of 1% in overall population size and urban population increases kerosene demand by 0.96% and 0.59%, respectively.

6.14 Biomass demand

The biomass analysed in this study is solid biomass like wood fuel, charcoal, etc., used by the majority of consumers in SSA, most are traded in the informal sector of the economies under study. Firstly, we present the correlation matrix of the variables. Secondly, the charts of the income and biomass demand variables are plotted. Lastly, the results of the fixed effects and random effects models for biomass demand in SSA are presented in Table 6.19.

Table 6.19: Correlation matrix of the variables

	lnIncome	lnPopulation	lnUrban	lnEconomy	lnBiomass
lnIncome	1.000				
lnPopulation	-0.320	1.000			
lnUrban	-0.147	-0.121	1.000		
lnEconomy	0.259	-0.053	0.183	1.000	
lnBiomass	-0.279	0.936	-0.007	0.029	1.000

As can be seen from Table 6.19 above, we do not have a multicollinearity problem among the variables analysed. Although evidence of high collinearity is found between population and biomass (0.936), nothing is done because the model has a high R-squared of 0.880, which indicates good fit.

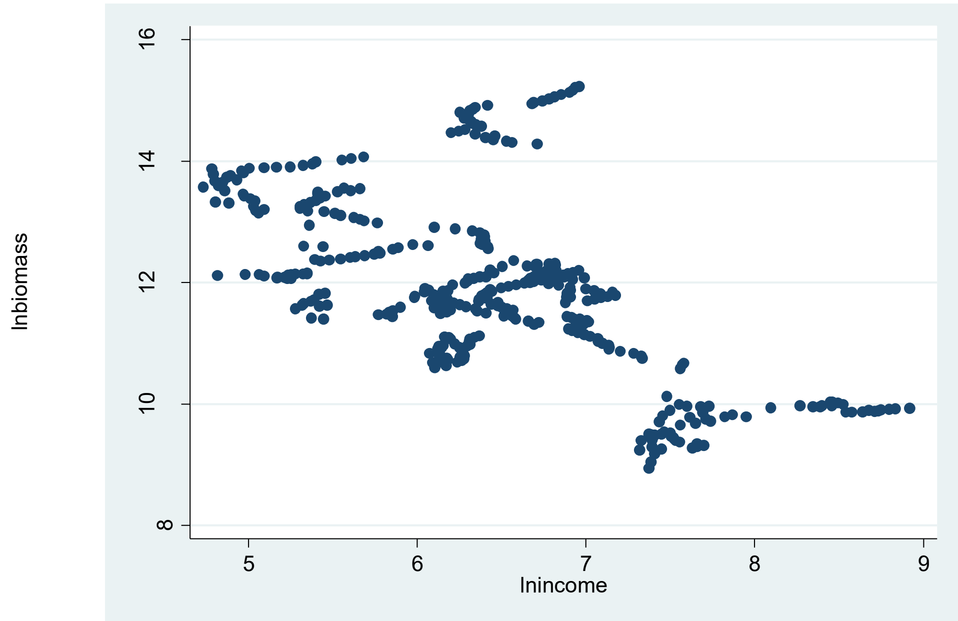

Figure 6.18: Scatter plot of biomass demand and income

Interesting from Figure 6.18 above, we can see that a negative relationship exists between biomass demand and income. This is reasonable because the higher the income of the consumer, the lower the amount of solid biomass consumed as it is an inefficient traditional form of energy.

Table 6.20: Estimation results for biomass demand models

Dependent variable: log of biomass	Panel A: Fixed Effects		Panel B: Random Effects	
Explanatory variables	Coeff	t-stat	Coeff	t-stat
lnIncome	-0.036	-1.30	-0.036	-1.32
lnUrbanisation	0.087^{***}	3.91	0.089^{***}	4.03
lnEconomic structure	-0.028	-1.10	-0.021	-0.99
lnPopulation	0.886^{***}	32.27	0.896^{***}	33.10
Constant	-2.773^{***}	-5.52	-2.931^{***}	-5.49
F test for model significant				
F-statistic for fixed effects and Wald statistic for random effects and PW	287.49		34.00	
P-value>F	0.000		0.000	
Observations	402		402	
Groups	12		12	
Model goodness-of-fit				
R-squared				
Within	0.748		0.748	
Between	0.880		0.880	
Overall	0.873		0.874	
Autocorrelation test (Wooldridge test)			115.372^{***}	
Heteroscedasticity test (Wald test)	68311.36^{***}			
Breusch- Pagan selection test	5438.85^{***}			
Breusch- Pagan test decision	Random effects model over OLS			
Hausman Test	Fixed vs Random effects			
Test statistic	6.25			
p-value> Test statistic	0.1811			
Decision	Random effects model			

Note: *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$

From Table 6.20, we will be using the random effects model chosen by the Hausman test. The selection of the model is likely to be due to the reason given above (see Section 6.10 above) linked to the prevalence of informal markets in the biomass industry.

The random effects model shows that urbanisation and population are the key variables influencing biomass demand in SSA. The negative income elasticity is in line with our *a priori* expectations that as consumers' income increases they will use modern energy sources. From this model, a 1% increase in consumer income reduces biomass demand by 0.04%. It is evident from the results that the wide prevalence of biomass use in the SSA energy mix is due to the widespread poverty and low income level that are prevalent in these countries. However, a note of caution is due here because the elasticity is not statistically significant.

High population growth is also another reason responsible for the high use of solid biomass in the region. Specifically, the obtained elasticity indicates that a 1% increase in population leads to a 0.90% increase in biomass consumption. Also, many of the urban areas in the region have slums where most of the residents use biomass for cooking, in form of charcoal. This could explain the positive elasticity found for the urbanisation variable.

However, the negative relationship between biomass consumption and economic structure may be due to the fact that an increase in the service oriented sector will reduce the amount of biomass used in the economy, if those used by industries have reduced over the years. This should reduce the overall biomass consumption in the countries in the region.

6.15 Chapter summary

This chapter presented and discussed the main driving forces of aggregate energy demand in Sub-Saharan Africa, and then the disaggregated analysis by type of fuel. The latter is based on the factors that impact on the demand for petrol, LPG, diesel, electricity, biomass and kerosene in SSA, which are reported and explored in turn. Overall, the estimation results indicated that income, energy price, economic structure and degree of urbanisation are the main determinants of energy demand in Sub-Saharan Africa. All the identified factors have different impacts which support or contrast some of the existing findings in the literature. In general, also in the case of the disaggregated model for individual energy types, the results corroborate that income, urbanisation, economic structure and population are the key factors that drive the demand for the energy types analysed.

Chapter 7 : Further Discussion of the Significance and Contribution of the Results

7.1 Chapter overview

Using the results from the analysis presented in the previous chapter, this chapter discusses the findings further. The discussion of the significance and contribution of the findings connects to the results of previous empirical studies in areas related to the objectives of this study. The chapter is divided into different sections according to each of the energy type analysed. Nevertheless, it is worth noting that since no other published research has analysed both the driving forces of aggregate energy demand and disaggregated energy demand by fuel types in Sub-Saharan Africa in a single study, a direct comparison of the results with previous empirical work is not straightforward.

7.2 Discussion of the aggregate demand model

The main question in this study seeks to identify and analyse the factors that drive energy demand in the long run in SSA. Our results found support for the positive relationship between energy demand and income postulated theoretically. Similarly, the theoretically expected negative relationship between energy price and energy demand was also confirmed by the results. Hence, energy is a normal good. Both findings are consistent with the results presented in existing literature. For example, Al-Azzam and Hawdon (1999) had an income and price elasticity of 0.95 and -0.22 respectively in Jordan, Iwayemi *et al.* (2010) had an income elasticity of 0.66 and a price elasticity of -0.11 in Nigeria.

However, despite the fact that the cited studies also found a relatively inelastic price and income elasticity in their analysis, the value for the reported coefficients (elasticity) differ from our results (income elasticity of 0.10 and price elasticity of -0.46) probably due to the differences in the countries analysed and the use of the panel cointegration approach in the analysis, which is more reliable. Although, Iwayemi *et al.* (2010) employed cointegration in the analysis, it was in the time series context as only Nigeria was considered in the empirical work. The result is further corroborated by the findings reported by Amusa *et al.* (2009) in relation to South Africa, Kuma (2008) for Fiji, and De Vita *et al.* (2006) for Namibia.

All the studies highlighted so far, however, suffer from the fact that the impact of growth in urban population known as the degree of urbanisation and economic structure were not included in the models. This is one significant contribution of the present study since it has been accepted by many scholars that urbanisation is a key driver of energy demand (and hence energy consumption) in developing countries (Adom *et al.,* 2012; Mensah *et al.,* 2016).

However, the failure to identify urbanisation as one of the driving forces may be due to institutional differences in the countries analysed or the time period covered in the study. As it will be seen in the sub-sections that follow under specific energy type demand models, most of the factors included in our model were used in energy modelling by authors who studied energy demand in SSA, especially in the last five years (Adom *et al*., 2012; Adom and Bekoe, 2013; Adom, 2013; Mensah, 2014).

Although there is hardly any empirical study that has investigated the relationship between aggregate energy demand and urbanisation, some studies have analysed the impact of urbanisation on energy use. For instance, Poumanyvong and Kaneko (2010) investigated whether urbanisation leads to lower energy use and lower $C0_2$ emissions using a cross-country analysis. In contrast to our result, where a significant positive relationship was found between aggregate energy demand and urbanisation, they found that in low income countries urbanisation decreases energy use. The difference in findings may be due to three possible reasons. Firstly, the researchers have used a different data set, as the countries analysed in this study fall under lower income, lower middle and upper middle categories. Hence, considering that we have not separated them into these income categories we may not be able to compare the findings of the two studies directly. Secondly, the time period analysed in the two studies differs because we have considered the more recent period from 1980 to 2014 whereas Poumanyvong and Kaneko (2010) used an older dataset based on the sample period 1975-2005. Lastly, the econometric models employed are different (here in this part of the study, a panel cointegration approach was used).

The positive relationship found between urbanisation and aggregate energy demand is interesting. According to a report by the African Development Bank, the annual rate of urbanisation is about 3.5% with 32.8% of people in SSA living in cities (AfDB, 2012). Besides, most of the stock of vehicles, modern energy equipment and gadgets are in urban areas. Moreover, most houses in the urban areas are connected to the national grid which leads to an increase in energy use by the consumers who can now acquire more electrical gadgets as they move from rural to urban areas with increased access to electricity. It is, therefore, expected that an increase in urban population size will increase the overall amount of energy consumed, as found in our analysis.

Likewise, most empirical studies in the literature do not include the analysis of the relationship between aggregate energy demand and economic structure, but some studies have analysed the impact of structural changes on specific energy types consumption in individual SSA countries. The negative relationship reported between economic structure and aggregate energy demand is in contrast with the findings of Mensah *et al*. (2016). The authors found a positive relationship between economic structure and electricity consumption in Ghana. The difference in findings could be due to differences in countries analysed and the methodological approach.

7.3 Discussion of the electricity demand model

This study set out with the aim of estimating the coefficients of the identified driving forces of electricity demand in SSA. We found that the most significant determinants of electricity demand in the long run for the analysed countries in SSA are income, the degree of urbanisation and population. The positive relationship found between income and electricity demand is supported by findings of several empirical studies in the literature. These results are consistent with those obtained by Adom *et al.* (2012) who estimated the long run income elasticity of Ghana to be 1.59, Expo *et al.* (2011), who reported an income elasticity of electricity of 0.58 in Ghana, and De Vita *et al.* (2006) who reported 0.59 as the income elasticity of electricity in Namibia. Clearly, the income elasticity of electricity demand is inelastic and positive in the long run in SSA.

The explanation given for the results by most of the studies mentioned above is that as the income of consumers' increases, they are able to buy more gadgets and appliances which need electric power to function and eventually lead to an increase in the demand for electricity. This view is supported by the findings in this study and it further shows that as the standard of living improves, people will consume more electricity.

Surprisingly, the degree of urbanisation was found to have a negative relationship with electricity demand in SSA in our analysis. Although, this result differs from that reported by some previous published studies (see, e.g., Holtedahl and Joutz, 2004; Adom *et al.,* 2012), it is consistent with that of Adom and Bekoe (2013). Considering that most of the countries analysed are in the low income group, it has been suggested that the negative sign of the urbanisation coefficient may be due to 'urban compaction' (Poumanyvong and Keneko, 2010). According to the authors, the negative sign is likely to be due to the fact that most of the countries lack access to public infrastructure which may need more energy if the number of access areas increases. This explanation is in line with our findings, as most public services like rail lines, well equipped hospitals, uninterrupted power supply, to name but a few, which are obtainable in most big cities in the world, are not available in most of the urban cities in SSA.

The negative relationship between electricity demand and urbanisation found in this study may also be partly explained by the fact that most of the urban dwellers in SSA live in slums. World Bank data show that about 1 billion people in the world live in slums due to the high rate of urbanisation (World Bank, 2009). Considering that most of the slum dwellers may not be able to afford a legal electricity connection, some do so by connecting illegally through neighbours who are connected legally, while the rest simply rely on traditional means of energy like solid biomass. This situation may explain why 127 million - out of a total of 220 million urban population in the world - without access to electricity are in SSA, making the region the highest with urban population without electricity access globally (World Bank, 2010).

Therefore, an increase in urban population may put more pressure on the available electricity because of the prevalence of theft through illegal connection, which leads to more rationing and even a decline in the total amount of electricity available or consumed in urban areas.

The result of the positive relationship between electricity consumption and population size confirms the association between energy use and population growth. The elasticity shows a relatively elastic relationship. Every percentage increase in population, there will be 1.02% increase in the demand for electricity. This finding has many important implications for SSA countries because of the projected figure of 2.4 billion people in SSA by 2050 (PRB, 2013), and the low electrification rate of 32% according to the International Energy Agency (IEA, 2015).

The availability of electricity for the whole region will be a huge task because of the high amount of capital investment required, and the pace of providing electricity may not be able to meet up the high rate of population growth. According to the IEA, an extra $450 billion investment in the power sector is required to achieve a 100% access to electricity in the urban areas and to reduce power outages to 50% in the region (IEA, 2014). Another report by the World Bank asserts that the amount of power available per person in SSA has reduced in the last few decades because population growth has been higher than investments in generation capacity (World Bank, 2010).

This is a considerable challenge and the governments in SSA will need to encourage good family planning practices, otherwise, it may be impossible to provide the much needed power for the whole population in the region. Moreover, considering the important role played by the availability of regular power supply in the socioeconomic development of a country, the governments and policy makers must find a way to have a good balance between the rate of population growth and investments in important public infrastructure like electricity. The result of the electricity model presented is similar to Kebede *et al.* (2010) who also found a positive relationship between electricity demand and population in SSA.

7.4 Discussion of the diesel demand model

This study found that income, the degree of urbanisation and population size, are the main driving forces of diesel demand in SSA. The results show that as consumer income increases, there will be an increase in the amount of diesel consumed, *ceteris paribus*. This result is in line with demand theory. This finding is also in agreement with the results obtained by De Vita *et al.* (2006) for Namibia, and Abdullahi (2014) for Nigeria. Abdullahi (2014) suggested that the upward trend in diesel demand in Nigeria could be linked to manufacturing, telecommunication and high income household segments of the economy. This is the case for most of the countries in SSA, even in South Africa due to the failure of the government to upgrade or replace the installed power plants since the 1960s. The inadequate power supply has led to the wide use of diesel powered generating sets by most firms and households in the sectors mentioned earlier.

Contrary to *a priori* expectations, this study did not find a positive relationship between diesel demand and the degree of urbanisation. The observed negative relationship may be partly due to the deindustrialisation occurring in most SSA cities, due to poor infrastructure. If the number of manufacturing firms reduces, then the amount of diesel consumed in the urban areas will also be affected because most of the diesel consumed is used by firms' diesel generators, machineries, equipment and trucks.

In a special report on business in Africa by The Economist, it is stated that '*Africa lacks most of the things a successful manufacturing sector requires'* (The Economist, 2016, p.7). The special report featured a new tomato processing company developed by the Dangote Group in Nigeria as a case study. They reported that 400 litres of diesel are used per hour in the plant to generate power. The high cost of power generation in the featured case perhaps sheds some light to the cause behind the high rate of deindustrialisation in the region. In other words, high overheads associated with providing power using diesel generators to operate machineries may explain the decline in diesel demand due to the inability of companies to sustain operations, especially those companies situated in urban areas. This is also evident from the statistics of the biggest economy in the region that show that 36% of production costs relates to generation of power privately by firms (Okafor, 2008).

It is interesting to note that a positive relationship was found between diesel demand and population. This result may be explained by the fact that as consumers' income increases and new firms enter the market, they will acquire diesel generators, machineries and trucks for transport. Despite the high cost associated with private generation of power (using generators), the high population has provided viable investment opportunities because of the large market size. Furthermore, the emergence of a new middle class in most countries has also enabled more people to acquire more appliances and gadgets which use electricity mostly provided by generators. This is partly due to the fact that the power generation capacity has not been able to meet up with the rate of population growth in SSA, as stated in the previous section. Moreover, diesel is an important fuel used in the transport sector of the economies, and an increase in population leads to an increase in the number of people needing transportation services.

7.5 Discussion of the LPG demand model

The results of this study show that income and population are the main determinants of LPG demand in SSA. The result of a positive income elasticity is consistent with that reported in earlier studies. Earlier empirical studies with similar findings include Abdullahi (2014), Mensah (2014), Ackah (2014), Akinboade *et al.* (2008), Alves and Bueno (2003), with reported income elasticities of 0.64, 0.45, 1.95, 0.36 and 0.12, respectively.

A possible explanation for the positive income elasticity may be that as income increases, consumers move away from traditional energy types like solid biomass and kerosene to a more modern energy type. Another possible explanation for the relatively elastic income elasticity of 1.10 obtained in our analysis may be due to the fact that any increase in LPG price will lead to a switch to alternative, cheaper options like traditional solid biomass and kerosene. The result is consistent with the 'energy ladder' hypothesis, which predicts that as the income of households increases, they move up the energy ladder by stepping up from the use of traditional energy to the use of modern energy sources like LPG (Mensah *et al.,* 2016).

Likewise, the positive relationship found between LPG demand and population size seems plausible because with the emergence of a new middle income class in SSA, most consumers are able to afford the use LPG in homes for cooking. Also, an increase in population would mean an increase in market size for firms that use LPG in their production activities like those in the hospitality sector.

7.6 Discussion of the petrol demand model

One of the most significant contributions of the present study is the disaggregation of the analysis by energy type so as to identify the specific response (elasticity) of each fuel type to changes in the associated explanatory variables. The estimation of the petrol model is quite revealing as discussed in Chapter 6, as more informed and focused policy recommendations can be made with regard to petrol demand management in SSA. The results of the elasticities as discussed in the previous chapter - were all significant and in line with our *a priori* expectations, except for the urbanisation variable, which recorded a negative sign.

It was found that the income price elasticity is positive, inelastic and statistically significant. This gives the government the opportunity to generate income from this market by either increasing the tax on petrol or removing the subsidy on the product so the money can be used for other capital projects, in countries where this policy is adopted. It is more prevalent in West Africa (Kojima and Matthews, 2010). This recommendation is proposed because the inelastic nature of the demand for petrol suggests that it is considered an essential product by consumers and an increase in price will result in only a modest reduction in demand.

Likewise, when consumers have a higher income they will not consume more than the needed amount but even if they need to pay more to meet the current consumption level, they are likely to do so because petrol is an essential good to them. Considering that most vehicles use petrol and even those without one, need it to power their power generating set due to the erratic power supply in the region.

However, the case for removal of the fossil fuel subsidy is not as straightforward because it will depend more on the individual country-context. This is in line with Bazilian and Onyeji (2013), who assert that the existing structure of reform and policy is important in determining how and if fuel subsidy or reform will be the way forward for a country. They argued further that despite the fact that this might be a country/context-specific question there are some similarities which are universal.

The consideration of the cost and benefits of a policy on fuel subsidy removal is important so as to prevent or reduce the negative impacts that could result from the removal (for example, see Wesseh and Lin, 2017, for discussion on Ghana). Some of the possible adverse effects include the reduction in modern energy use (if fuel becomes more unaffordable to the poor) and a higher cost of production to firms in energy-constrained countries, which would impede the establishment of small and medium-sized enterprises.

Despite the mentioned negative impacts, the benefits of removing the fuel subsidy are greater because such removal frees up capital which can be used for providing the much needed infrastructure, encourage energy efficiency due to the higher price charged, more private sector participation in the sector especially with building of refineries as higher prices in the market will make the investment viable to both local and foreign investors. Nevertheless, subsidy is well intentioned because its primary purpose by the government is mainly for wealth redistribution, to foster the development of energy-intensive industries, poverty alleviation through the promotion of industrialisation, amongst others.

However, the mentioned benefits are not really benefiting the poor who are the main target of such policy, but rather consumers on higher income (Bazilian and Onyeji, 2013; IEA, 2010; Victor 2009). Based on the ongoing discussion, it is apparent that before the fuel subsidy removal is implemented, there must be good reforms that will promote the availability of power in the country and other necessary energy infrastructure.

Our positive income elasticity for petrol demand is consistent with the findings of previous energy demand studies in the literature. These include Mensah *et al.* (2016) with an income elasticity of 1.32 for petrol demand in Ghana, Abdullahi (2014), who reported an income elasticity of 0.11 for Nigeria, Iwayemi *et al.* (2010) who reported an income elasticity of 0.75 for Nigeria, Akinboade *et al.* (2008) who reported an income elasticity of 0.36 for South Africa, and De Vita *et al.* (2006) who also found an elasticity of 1.27 for petrol demand in Namibia.

The negative link between petrol demand and urbanisation may be due to the fact that rural migrants slowly move from the use of solid biomass to modern energy. The switch is slow as rural migrants are low skilled workers, initially live in slums, unable to afford to buy a vehicle or even a small power generating set until after some time. This slow transition may explain the negative elasticity because some of the rural migrants may even become worse off in the city, and their standard of living will reduce due to the reduced disposable income. Most of the rural-urban migrants are likely to be farmers or unemployed youths in the rural areas; with the hope of a decent job and reasonable income immediately they get to the city.

However, their income may improve over time but the modernisation towards green and efficient energy sources will take time, before they are able to acquire petrol-using equipment will take even a longer time. This is supported by the statistics showing that SSA has the highest proportion of slum dwellers in Africa at 65% (AfDB, 2012). The result of the negative impact of urbanisation on petrol demand is in contrast with that of Mensah *et al.* (2016) who reported a positive link between petrol demand and urbanisation in Ghana. It seems possible that the differences in the results may be due to the fact that the present study is a cross-country analysis while that of Mensah *et al.* (2016) is solely a time series analysis of only one country.

We also found that the higher the industrial output in SSA, the higher the demand for petrol. This is as one would expect. Since there is inadequate power supply in the region, most commercial and residential consumers rely on the use of power generating sets to alleviate the inadequacy in supply. The two main fuels used in fueling the generating sets are petrol and diesel. This is evident, for example, by the high cost of 36% attributed to cost of private power generation by industries in Nigeria (Okafor, 2008, as cited in Bazilian and Onyeji, 2013).

Another logical explanation for this result is that an increase in industrial output will result in more transportation services. Since most of the countries in our sample do not have an efficient mass transit system in operation, they rely mostly on vehicles, vans, cars and trucks for transport and haulage. Therefore, an increase in industrial output will lead to an increase in petrol demand in SSA.

Lastly, our result supports the hypothesis that population growth increases energy use (consumption). The positive relationship found between population and petrol demand in this study is supported by one of the findings of Kebede *et al.* (2010). To be specific, the authors found a positive relationship between population and energy demand. The energy types included in their study are traditional energy, electricity and petroleum.

The analysis of petrol demand undertaken here, has extended our knowledge of how the driving forces of petrol demand impact on its consumption in the countries included in our sample.

7.7 Discussion of the kerosene demand model

The most obvious finding to emerge from the kerosene model analysis presented in Chapter 6 is that, income, degree of urbanisation and population, are the key drivers of kerosene demand in SSA. All the elasticities have the signs expected, and the results are supported by prior empirical studies in the literature. The positive income elasticity could be explained by the move towards more efficient energy sources like kerosene from traditional biomass like charcoal and firewood, as people's income increases. The relatively inelastic coefficient also suggests that as kerosene is the cheapest form of modern cooking fuel, consumers are less responsive to changes in price because it is an essential good to them. Studies reporting a similar result include Iwayemi *et al.* (2010) and Abdullahi (2014).

The analysed positive relationship between kerosene demand and the degree of urbanisation is likely to be due to the fact that when consumers move from rural areas to urban centres, they switch from traditional sources like firewood to kerosene cooking stoves. Even for slum dwellers the use of solid fuel like firewood may not be appealing because of the compacted style of living, and consumers may be able to afford small cooking stoves that use kerosene in the available living space. Moreover, with more awareness of the adverse health implications associated with the use of biomass sources for cooking in main cities, most households prefer to use kerosene stoves which are relatively more affordable in most SSA cities.

7.8 Discussion of the biomass demand model

As mentioned in the literature review, very few studies in the literature analysed the factors that determine biomass demand in SSA. Owen *et al.* (2013) is a notable exception. In an investigation into energy policies in SSA, Owen *et al.* (2013) found that biomass is seen by the policy makers as an inferior form of energy that promotes poverty and pollutes the environment. This may explain why most of the reviewed studies did not include biomass in their analysis of energy demand.

However, the present study includes solid biomass as one of the main energy types analysed, because it accounts for over 80% of the total energy consumed in the region. According to the IEA, four out of five people rely on fuelwood for cooking in SSA (IEA, 2014). Hence, it is an important energy type that should be included in the analysis of the driving forces of energy demand in the region. The results of this study indicate that urbanisation and population are the significant factors that drive biomass demand.

As discussed earlier, the rapid rate of urbanisation and population growth has not been matched by increased availability of energy in SSA. Coupled with the widespread poverty especially in rural areas, most of the low income earners rely on the use of solid biomass for cooking. A low income level is prevalent in most SSA cities as well as rural areas. Hence, a rapid population growth rate leads to an increase in the number of people using solid biomass for cooking. The study by Rahut *et al.* (2017) also found that poorer households tend to use solid biomass, kerosene and batteries for lighting.

Considering that both elasticities are relatively inelastic, it is reasonable to acknowledge that people on low income consuming biomass are less responsive to changes in prices because it is the cheapest form of energy available. Moreover, most of the people in rural areas collect wood fuel directly from the forest or farms. This finding is corroborated by Kebede *et al.* (2010) who also reported a positive relationship between energy consumption (traditional fuelwood) and population in SSA. These results are also in agreement with Mensah *et al.* (2016) who showed that income and urbanisation are the main drivers of biomass demand in Ghana. Mensah *et al.* (2016) showed that income is a significant determinant of biomass demand in Ghana. This differs from the findings presented in this study. Income was found to have a negative albeit insignificant relationship with biomass demand in SSA. The negative sign is as expected. Consumers with more income would demand less biomass and move towards the use of modern energy like kerosene and LPG. However, a possible explanation for the statistically insignificant coefficient may be linked to this being a cross-country study and our analysis shows what is happening across the countries analysed, while the study by Mensah *et al.* (2016) is a single country study.

In general, therefore, it seems that the high proportion of urban poor, a high level of unemployment and the shortage in modern energy types like LPG and electricity, may have led to the use of different forms of biomass among the low income earners who are unable to afford the use of generators or modern cooking stoves.

7.9 Chapter summary

This is the first study to investigate and analyse the determinants of both aggregate and disaggregated energy demand in SSA. The findings from this study make several contributions to the current literature. Firstly, the reported elasticities for income, price, economic structure, degree of urbanisation and population, for the different energy types analysed, extend our knowledge of the determinants energy demand in SSA. Secondly, the analysed elasticities can be used to guide informed energy demand management policies in the region. Thirdly, both public and private investors in the energy sector can use the findings of the study to determine if investment in this sector in SSA would be viable. Lastly, the driving forces identified will assist our understanding of the role they play in energy demand. Also, the energy model employed can be used to analyse the determinants of energy demand elsewhere in the developing world.

The results of the analysed energy demand models were highlighted and discussed one after the other in light of the findings of existing literature. Some similarities and a few differences were found in comparison to the results of existing studies and logical explanations to explain such differences presented.

Clearly, there is need for the governments and policy makers to harness the abundant energy resources (see Chapter 2) in providing the much needed energy in the region so as to foster economic and social development.

Chapter 8 : Conclusions

8.1 Chapter overview

An efficient demand management is needed to boost the socio-economic development of a nation. All countries in Sub-Saharan Africa (SSA) have been unable to provide the energy needed by consumers to achieve economic prosperity and improve the standard of living. Considering that this requires huge capital investments and careful planning, sound evidence-based policies are required to make well informed decisions. In the next section of this chapter, the summary of the research findings linked to the specific objectives of the PhD study are presented. Some policy implications are drawn from the findings and presented in the next section, followed by a statement of the study's main contribution to knowledge. The chapter ends by acknowledging some limitations of the research and by highlighting areas for future research.

8.2 Summary of research findings

Although a number of mostly single-country studies have investigated the determinants of energy demand in some African countries, no single study has examined the elasticities of energy demand at aggregate level and by energy type using the latest panel data model techniques for a large proportion of representative SSA countries. The primary aim of this thesis is to fill this gap in the literature by identifying and analysing empirically the main factors that drive energy demand in Sub-Saharan Africa. The impact of income, price, urbanisation, economic structure and population on aggregate and disaggregated (fuel type) energy demand in the SSA region is explored in the thesis. The specific objectives of the research were:

(i) To provide a full and comprehensive analysis of the energy sources in the region;
(ii) To critically review both the theoretical and empirical literature on energy demand (energy consumption) in developing regions including SSA;

(iii) To provide a review of the aggregate and disaggregated pattern of energy demand literature in the region;
(iv) To provide preliminary conclusions on the main energy issues and prospects affecting the region;
(v) To develop a comprehensive econometric model for the analysis of a cross-country aggregate and disaggregated (energy type) energy demand function for SSA to estimate elasticities of energy demand to changes in the explanatory variables;
(vi) To make an original contribution to the existing body of knowledge on energy demand in SSA and draw out relevant policy implications in light of the research findings.

As presented in the previous chapters of the thesis, all the stated objectives have been met. The sections below explain how each objective was met.

8.2.1 (Objective i) Analysing the Sub-Saharan Africa energy sources

An account of the energy sources in SSA is provided in chapter 2.

Chapter 2 is inevitably descriptive yet most informative in nature. It provides a detailed account of the different energy sources available in Sub-Saharan Africa. Both the renewable and the non-renewable energy sources available in the region were discussed in detail using relevant statistics, to provide greater insights about the vast energy sources in the region. The chapter also explored the specific characteristics of the energy consumption pattern in the SSA region, to guide the selection of a suitable model for the analysis. SSA has abundant fossil fuels and uranium resources located in different countries across the region. There is a huge hydro potential in the Central African region, large oil and gas reserves in the West, a high renewable potential (coastal area vast wind energy and the eastern part huge geothermal sources) and the large coal deposits in the southern parts. Angola and Nigeria are the largest oil producers in the region, and are also home to vast natural gas resources. Namibia, South Africa and the Republic of Niger are among the top ten uranium reserve holders globally.

From the analysis presented in Chapter 2, it is evident that interregional co-operation of trade and supply of energy is the way forward for the region considering the uneven distribution of the energy resources, capital constraints and the low income level in the region. However, good transmission and connection infrastructure would have to be built to facilitate interregional trade and supply mechanisms.

8.2.2 (Objectives ii and iii) Review of the energy demand literature

Chapter 3 is devoted to a critical review of the literature on energy consumption at the aggregate and disaggregated level, energy efficiency and carbon emission - mostly in developing countries. This macro level study is underpinned by the microeconomics-based neo-classical theory of consumers' utility optimising behavior. Existing evidence shows that sufficient energy is needed to unlock the economic and development potential of SSA. From the literature review, it could be gauged that none of the prior studies have analysed the determinants of both aggregate and disaggregated energy demand in the region. Such analysis requires both the identification of the key determinants of energy demand and the use of reliable long macroeconomic data.

Chapter 4 provides an overview of the panel data methods employed in this study. The chapter explored the definitions, derivations and interpretations of the panel methods used in the research. The SSA compiled dataset is presented in Chapter 5, alongside the sources of the data and their measurement. Considering that the secondary data are sourced from publicly available and widely used reliable energy and economic databases, valid results are obtained from the analysis done using STATA 13.

8.2.3 (Objectives iv and v) Energy demand model and estimation results

Chapter 6 presents the analysis and results of the econometric models employed for the analysis of the aggregate and disaggregated (energy types) energy demand in SSA. The main variables used to analyse the demand for energy were chosen based on the literature review, that is, price, income, urbanisation, economic structure and population. The analysis of the aggregate energy demand revealed that income (of the consumer), price and urbanisation are the main factors that drive the demand for energy in the region. When consumers earn more, they acquire energy using gadgets and appliances which use energy. This increases the total amount of energy consumed. Similarly, if the price of energy increases, the amount of energy consumed reduces. Our findings are in line with demand theory and the estimated elasticities are found to be relatively inelastic. Furthermore, with an increase in urban population, more people have access to modern energy and acquire more energy appliances and gadgets, which lead to an increase in the total amount of energy demanded (consumed).

Energy types including electricity, diesel, liquefied petroleum gas (LPG), petrol, kerosene and solid biomass are also analysed. The results are presented in Chapter 6. In all the analysed models, population is the predominant factor behind the increase in demand because it has the largest elasticity. As population increases, there is more demand for energy (by the increasing population). Electricity and petrol have the highest population elasticity coefficient, recording a unit elasticity. For every percentage increase in population, the demand for electricity and petrol increases by 1.02% and 1.12%, respectively, *ceteris paribus*. It was also found that for all energy types, except solid biomass (as to be expected), income is one of the main determinants in the long run in SSA. As the income of consumers increases, they move away from traditional sources of energy which dominate the region energy mix, toward more modern energy types. This tendency also reduces the time wasted on collecting solid biomass (firewood), health-related issues and economic drawbacks associated with the use of solid biomass for cooking in the residential sector.

Significantly, the study also found that urbanisation too is a significant factor behind the demand of the analysed energy types, except LPG. With the increased trend in the number of people living in urban areas in SSA, they can access more modern energy equipment in cities, which has a positive impact on the total amount of energy consumed. Also, considering that most of the appliances are not energy efficient, due to the importation of used goods from developed countries and other developing countries, more energy is consumed as the stock of appliances increases. There is a need for good energy efficiency regulations with regard to imported goods, to ensure that the available energy is used in an efficient way. This would also reduce the large energy supply-demand deficit in the region. All in all, energy demand in SSA will continue to increase in the long run alongside an increase in supply. However, increased availability of (access to) energy is needed to foster socio-economic development in SSA. All the results are consistent with some of the existing findings in the literature pertaining to the experience of specific developing countries. Further detailed discussion of the importance of the empirical findings is provided in Chapter 7.

8.3 Policy recommendations

This section relates to **objectives iv and vi**. Policy implications are drawn from the reported energy demand elasticities estimated and discussed in Chapters 6 and 7.

The study has identified and investigated the long run effects of the driving forces of aggregate and disaggregated (energy types) energy demand in SSA. The findings of the study are not only of interest to academics but also policymakers and other stakeholders in the region. With the increasing trend in population growth, income and urbanisation in the region, there is a need for the relevant governments and policymakers to provide an efficient mechanism of how the growing demand of energy will be met with adequate supply. The following recommendations are made based on the research findings:

1. A key finding was the positive impact of urbanisation on the aggregate energy demand, hence, a critical policy implication that follows is to reduce or curb large rural-urban migration. In particular, the following steps are recommended:
    - Economic growth should be spread out across all regions in the countries by creating employment opportunities in small towns, so that young people do not have to migrate and live in slums in urban areas due to economic reasons.
    - Small scale renewable projects like small roof top solar panels should be introduced and promoted in off-grid areas to cater for

small businesses energy need, and to improve the overall standard of living.

- Policies that will promote the affordability of small scale off-grid energy sources should be promoted by the governments in the region. This can be achieved, for example, by providing low interest soft loans for businesses in the rural area that need power to function or operate well.
- Energy access should also be scaled up by connecting rural areas to the national grid in order to attract more manufacturing firms and industries to provide the needed employment in these areas.
- Governments should subsidise the cost of charcoal cooking stoves and make them more available in rural communities so as to reduce further the time used in collecting solid fuel from the forest by girls and women.

2. Another key result shows that an increase in income will lead to an increase in energy use. The implication of the result is that there will be an increase in energy use across the region. Therefore, the policy recommendation from the result is to use the available energy efficiently. In order to achieve this, the following steps are recommended:
    - Energy efficiency guidelines and awareness should be made available with regard to the acceptable specification of appliances and gadgets to be imported, sold in markets and made available to consumers.
    - Government agencies in charge of the inspection of both used and new imported goods should ensure that they are energy efficient appliances and gadgets, when brought into SSA countries.

3. Considering GHG emissions associated with energy generation from fossil fuels, use of renewable energy sources to produce green (or cleaner) energy should be promoted. Where fossil fuel is still used, carbon pricing should be introduced to make energy firms more accountable.

4. Reliable and sustainable energy systems can be ascertained through proper electricity metering systems, which take into account differences in income across regions in the country. The billing strategy would ensure that enough revenue is generated and will also provide funds for future investments. Such a strategy is also likely to attract more private sector participation, as it would make investment in the sector more appealing and viable.

5. Another important finding (from the review of the literature) is that regional integration and cooperation of trade and energy generation should be promoted. To achieve this, the following steps are recommended:
   - More interregional pipelines should be built so as to provide more gas and other fuels for households, power plants and industries in SSA countries.
   - Effective energy power pools should be created so as to encourage investments in countries where abundant resources are available. This will encourage more private sector participation. For instance, the hydropower potential in the Democratic Republic of Congo is only attractive if the generated power can be sold to other countries in the region because the generated electricity will exceed domestic demand.
   - Build more interconnecting and transmission infrastructure to reduce the capital constraints and pressure on individual countries, to achieve the common goal of providing energy for consumers.

8.4 Contribution to knowledge

This research makes several, significant contributions to knowledge (objective vi).

First, a thorough critical reading and synthesis of the existing studies on energy demand in developing countries was conducted and used to identify the main factors that drive energy demand in the region. The identified factors are employed in the models used for the analysis. This research clearly shows the current dynamics of the energy situation in Sub-Saharan Africa (SSA), which future studies examining the determinants of energy demand in the region could also employ in their analysis.

Second, the study analysed and provided the first up-to-date empirical evidence of the main determinants of both aggregate and disaggregated (fuel type) energy demand in Sub-Saharan Africa using state-of-the-art panel data econometric techniques. This enabled the author to analyse the impact of the determinants of energy demand on aggregated energy demand and each of the analysed energy type. Accordingly, the thesis made a significant contribution to our existing knowledge by showing that the identified determinants have different impacts on the energy types analysed, through the estimated coefficients.

Third, the coefficients of the analysed factors have been used to suggest evidence-based implications and policy recommendations for ways to better manage and plan energy demand, in order to meet both the present and future energy need of the consumers across the SSA region. Knowledge of the signs of the estimated coefficients will also be beneficial to investors in SSA, shedding further light on the present and likely future trend of energy consumption across the region.

To sum up, the overall, original contribution of this thesis, can be highlighted by answering the question, 'what do we know now that we did not know before as a result of this thesis?' Theoretically, the estimations have confirmed that the law of demand is a suitable theoretical framework for energy demand. Energy consumption responds to price, income and other economic determinants. In contrast to some voices in the literature suggesting that energy is a luxury good, the study found that in the context of SSA countries energy behaves as a normal good, therefore responding positively to increases in income (with an aggregate elasticity of 0.10) and negatively to decreases in price (with an aggregate price elasticity of -0.46). This study also examined two variables that have rarely been examined in the energy economics literature, namely: the degree of urbanisation and economic structure. Thus, the findings from the study increase our knowledge on the impact of other economic factors on energy consumption in Sub-Saharan African.

8.5 Limitations of the study

This section explicitly acknowledges some of the limitations of the research. First, the study is limited by the lack of available data for all the countries in SSA, and also the lack of a price variable for the energy types analysed. Only 16 countries out of 47 SSA countries could be analysed in the aggregate energy demand analysis, while the sample for the disaggregated energy demand analysis by energy types, is based on 12 countries only. Nevertheless, the analysis provides a starting point by shedding light on what drives energy demand in the region, but of course the results may not be taken to be fully representative of all the individual countries making up the region. However, considering that most of the SSA countries have a similar energy situation and share a similar institutional structure, our findings certainly provide a true reflection of all the SSA countries analysed and possibly a rough snapshot of the key drivers of energy demand in SSA.

Another limitation pertains to the analysis of the existence of structural breaks which in this study was only undertaken through visual inspection of the plots of the evolution of the series. However, a more sophisticated analysis may be called for thorough appropriate panel unit root tests that account for one or more structural breaks in the series, in order to ascertain with greater certainty whether or not there are such breaks in the variables.

Similarly, the author also assumed a linear relationship among the variables analysed in the models. Although this assumption is not uncommon in the empirical literature, it may not be warranted. This potential extension to the analysis provides another profitable avenue for future research to model and test for non-linearities in both the time properties of the individual series and in the cointegrating relationship itself. That said, modelling and estimation techniques of nonlinearities in a context where the variables exhibit a mixed order of integration are still in their infancy. Moreover, the reliability of such techniques is still subject to debate, as recently shown in the replication study conducted by De Vita and Trachanas (2016).

In this PhD study, no test was conducted for reverse causality. That is, no (Granger) causality test was carried out to verify whether energy consumption causes growth, though the analysis demonstrated that economic growth causes energy consumption. It must be noted, however, that the testing of bidirectional causality was beyond the scope of the study, as the aim was solely to analyse the determinants of energy demand.

Finally, primary data could have been collected through interviews with various stakeholders, to gauge their perspective. This was excluded in the initial planning of the study due to time, cost and the risky travelling arrangements that attempting to obtain primary data across countries in SSA would have entailed. Moreover, this is an econometric study and the use of long macroeconomic data is preferable.

8.6 Avenues for future research

This research has made a significant contribution to knowledge by unveiling the key factors that impact on energy demand in SSA. The elasticities of the identified factors on energy demand have also been estimated. Yet, despite the inherent merits of this contribution, a number of open ends remain. This means that this research provides, additionally, a good starting point for further research on this important topic. Other researchers might wish to pick up from here to answer any questions that remain lingering on the sideline since they were beyond the scope of the study. Preferably, an effort should be made to explore how much energy could be made available from energy efficiency implementation in the region. This will give information about how much of the needed supply can be achieved from good energy efficiency programs in the region.

Another area for future study could be on the long run economic impact of the fuel subsidy removal in the face of falling commodity prices. Effort should be made to collect price data for the energy types in the region, to formulate effective energy price policies in SSA. Considering the ongoing debate about the removal of the fuel subsidy in some of the countries, analysing the demand with the inclusion of price data will provide more robust evidence about what might be the best policy in this area for such countries. This is very important especially with the plunge in commodity prices, as the fund used to subsidise fuel can be channelled to providing the needed infrastructure in the region. Otherwise, governments may continue the fuel subsidy regime if it will be economically beneficial in the long run for the masses. Further research has the potential to provide good evidence as to what might be the best option.

More broadly, research is also needed to forecast the future aggregate and disaggregated energy demand in the region. These forecasts would aid policymakers and inform investors as to what exactly to work towards, in terms of the exact amount of energy to be provided, and the level of investment that would be required to meet the expected demand, in addition to the information provided by existing research (such as in this current study) on what drives demand. Some existing studies have also made an attempt to forecast the energy required at individual country level, but having a more reliable cross-country forecast will take the analysis and associated debate a step forward.

The mentioned areas for future research do not reduce the significance of the contribution to knowledge made by the present study. This is because for the first time, a single study has analysed the driving forces behind both aggregate and disaggregated energy demand in SSA. Exploring the recommended areas for future research will nevertheless shed more light to ways of reducing the energy poverty in SSA.

References

Abdullahi, A. (2014) 'Modelling petroleum product demand in Nigeria using structural time series model (STSM) approach', *International Journal of Energy Economics and Policy*, pp.427-441.

Ackah, I. (2014) 'Determinants of natural gas demand in Ghana', *OPEC Energy Review*, pp.272-295.

Ackah, I. and Renatas, K. (2015) 'Green growth in oil producing African countries: A panel data analysis of renewable energy demand', *Renewable and Sustainable Energy Reviews,* 50, pp.1157-1166.

Adom, P. (2013) 'Time varying analysis of aggregate electricity demand in Ghana: a rolling analysis', *OPEC Energy Review*, pp.63-80.

Adom, P. and Bekoe, W. (2013) 'Modelling Electricity demand in Ghana revisited: The role of policy regime changes', *Energy Policy*, pp.42- 50.

Adom, P. and Kwakwa, P. (2014) 'Effects of changing trade structure and technical characteristics of the manufacturing sector on energy intensity in Ghana', *Renewable and Sustainable Energy Reviews,* 35, pp.475-483.

Adom, P., Bekoe, W and Akoena, S.K. (2012) 'Modelling aggregate domestic electricity demand in Ghana: an autoregressive distributed lag bounds cointegration approach', *Energy Policy*, 42, pp. 530–537

Adom, P., Bekoe, W., and Akoena, S. K. (2012) 'Modelling aggregate domestic electricity demand in Ghana: An autoregressive distributed lag bounds cointegration approach', *Energy Policy*, pp.530-537.

Adom, P., Bekoe, W., Amuakwa-Mensah, F., Mensah, J. and Botchway, E. (2012) 'Carbon dioxide emissions, economic growth, industrial structure, and technical efficiency: Empirical evidence from Ghana, Senegal, and Morocco on the causal dynamics', *Energy,* 47, pp.314- 325.

Afdb (2012) *Urbanisation in Africa.* Available at: http://www.afdb.org/en/blogs/afdb-championing-inclusive-growth-across-africa/post/urbanisation-in-africa-10143/ [Accessed 7th April 2016].

African Development Bank (2004). Strategic Plan for Wind Energy Development in Africa.

Akarca, A. and Long, T. (1980) 'On the relationship between energy and GDP: a reexamination', *Journal of Energy Development,* 5, pp.326- 331.

Akinboade, O., Ziramba, E. and Kumo, W. (2008) 'The demand for gasoline in South Africa: An empirical analysis using co-integration techniques', *Energy Economics*, pp.3222- 3229.

Akinlo, A. E. (2008) 'Energy consumption and economic growth: evidence from 11 African Countries', *Energy Economics*, 30, pp.2391-2400.

Alagidede, P., Adu, G. and Frimpong, P.B. (2014) 'The effect of climate change on Economic Growth, Evidence from Sub-Saharan Africa', WIDER working paper 017.

Al- faris, A. (1992) 'Income and price elasticity of gasoline demand in Organization of Arab Petroleum Exporting Countries', *Journal of Energy and Development*, pp.209-223.

Al-Azzam, A., and Hawdon, D. (1999) 'Estimating the demand for energy in Jordan: A Stock-Watson dynamic OLS(DOLS) approach', *Surrey Energy Economics Centre, Department of Economics, University of Surrey*, pp.1-17.

Al-Iriani, M. (2006) 'Energy- GDP relationship revisited: An example from GCC countries using panel causality', *Energy Policy*, pp.3342-3350.

Aliyu, A., Ramli, A. and Saleh, M. (2013) 'Nigeria electricity crisis: Power generation capacity expansion and environmental ramifications', *Energy*, pp.354-367.

Al-Kuwari, D. (2007) *Determinants of dividend payout ratio of companies listed on emerging stock exchanges: The case of the Gulf Cooperation Council (GCC) countries.* PhD thesis. Cardiff University.

Allcott, H. and Greenhouse, M. (2012) 'Is there an energy efficiency gap?', *NBER Working Paper Series,* pp.1-38.

Alter, N. and Syed, S. (2011) 'An Empirical analysis of electricity demand in Pakistan', *International Journal of Energy Economics and Policy*, pp.116-139.

Alves, D. C. and Bueno, R. D. (2003) 'Short- run, long-run and cross elasticities of gasoline demand in Brazil', *Energy Economics*, pp.191-199.

Amusa, H., Amusa, K., and Mabugu, R. (2009) 'Aggregate demand for electricity in South Africa: An analysis using the bounds testing approach to cointegration', *Energy Policy*, pp.4167- 4175.

Ang, J. (2008) 'Economic development, pollutant emissions and energy consumption in Malaysia', *Journal of Policy Modeling,* 30, pp.271-278.

Aqeel, A. and Butt, M. S. (2001) 'The relationship between energy consumption and economic growth in Pakistan', *Asia Pacific Development Journal,* 8(2), pp.101-110.

Aranda, A., Ferreira, G., Mainar-Toledo, M. D., Scarpellini, S. and Sastresa, E. L. (2012) 'Multiple regression models to predict the annual energy consumption in the Spanish banking sector', *Energy and Buildings, 49*, pp.380-387.

Asteriou, D. and Hall, S. (2011) *Applied Econometrics.* Second edn. China: Palgrave Macmillan.

Atakhanova, Z. and Howie, P. (2007) 'Electricity Demand in Kazakhstan', *Energy Policy*, pp.3729-3743.

Baltagi. (2008) *Econometric analysis of panel data.* Fourth edn. USA: John Wiley and Sons.

Baltagi, B. H. and Kao, C. (2000) 'Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey', *Center for Policy Research,* Paper 136.

Bazilian, M. and Onyeji, I. (2013) 'Fossil fuel subsidy removal and inadequate public power supply: Implications for businesses', *Energy Policy,* 45, pp.1-5.

BGR (2013) *Energy Resources 2013, Reserves, Resources, Availability, Tables*, Hannover, Germany.

Bhattacharyya, S. and Timilsina, R. (2009) 'Energy demand models for policy formulation: A comparative study of energy demand models', *The world bank development research group, environment and energy team*.

Bhattacharyya, S. and Timilsina, R. (2010) 'Modelling energy demand of developing countries: Are the specific features adequately captured?', *Energy Policy*, pp.1979-1990.

Bianco, V., Manca, O. and Nardini. (2009) 'Electricity consumption forecasting in Italy using linear regression models', *Energy*, pp.1413-1421.

BIldirici, M. E. (2013) 'The analysis of relationship between economic growth and electricity consumption in Africa by ARDL method', *Energy Economics Letters,* pp. 1-14.

Bispham, F. (2005) *Estimation and inference with nonstationary panel data.* PhD thesis. University of Hull.

Blazquez, L., Boogen, N. and Filippini, M (2013) 'Residential electricity demand in Spain: New empirical evidence using aggregate data', *Energy Economics*, pp.648- 657.

Breitung, J. (2000) 'The local power of some unit root tests for panel data', *Advances in Econometrics,* 15, pp.161-177.

Breitwieser, A. and Wick, K. (2015) 'What we miss by missing data: Aid effectiveness revisited', *World development,*78, pp.554-571.

Brenton, P. (1997) Estimates of the demand for energy using cross-country consumption data. *Applied Economics, 29*, pp.851-859.

Brew- Hammond, A. (2007) 'Challenges to increasing access to modern energy services in Africa', *proceedings of Forum of Energy Ministers of Africa Conference on Energy Security and Sustainability,* Maputo, Mozambique.

Broeck, J., Cunningham, S. A., Eeckels, R. and Herbst, K. (2005) ‘Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities’, *Plos Medicine,* 2(10), pp.0966-0970.

Breusch, T. S. and Pagan, A. R. (1979) A Simple Test for Heteroskedasticity and Random Coefficient Variation’, *Econometrica.* 47, pp. 1287–1294.

Byrne, D. (2002) *Interpreting Quantitative data.* London: SAGE Publications Ltd.

Cedigaz (2013) The 2013 Natural Gas Year in Review, *Institut Francais du Petrole*, Rueil-Malmaison, France.

Chakravarty, S. R. (2002). *Microeconomics.* Allied Publishers.

Changhong, C., Biagyan, W., Qingyan, F., Green , C. and Streets, D. (2006) ‘Reductions in emissions of local air pollutants and co-benefits of Chinese energy policy; a Shanghai case study’, *Energy Policy, 4*, pp.754-762.

Chen, C., Chen, B., Green, C. and Wu, C. (2002) ‘Benefits of expanded use of natural gas for pollutant reduction and health improvements in Shanghai’, *Sinosphere, 5*, pp.58-64.

Cheng, B. and Lai, T. W. (1997) ‘An investigation of co-integration and causality between energy consumption and economic activity in Taiwan’, *Energy Economics, 19*, pp.435-444.

Choi, I. (2001) ‘Unit roots test for panel data’, *Journal of International Money and Finance,* 20, pp.249-272.

Chontanawat, J., Hunt, L.C. and Pierse, R. (2006) ‘Causality between energy consumption and GDP: evidence from 30 OECD and 78 non-OECD countries’, *Surrey Energy Economics Discussion Paper Series*, 113.

Chudik, A. and Pesaran, M. H. (2013) Large Panel Data Models with Cross-Sectional Dependence: A Survey. *Federal Reserve Bank of Dallas*, August, pp.1-54.

Coghlan (2010) *Welcome to a Little Book of R for Time Series.* Available at: http://a-little-book-of-r-for-time-series.readthedocs.org/en/latest/#acknowledgements [Accessed 20 September 2015].

Collings, J. (2002) *Mind over matter: The Sasol story: a half-century of technological innovation*, Sasol.

Collins, L., Schafer, J.L. and Kam, C. M. (2001) 'A comparison of inclusive and restrictive strategies in modern missing-data procedures', *psychological methods,* 6, pp.330-351.

Culver, S. and Papell, D. H. (1997) 'Is there a unit root in the inflation rate? Evidence from sequential break and panel data model', *Journal of Applied Econometrics,* 12, pp.435-444.

Dale, A., Arbor, S. and Procter, M. (1988) *Doing secondary analysis.* London, UK: Unwin Hyman.

Dasappa, S. (2011) 'Potential of biomass energy for electricity generation in Sub-Saharan Africa', *Energy for Sustainable Development,*15, pp.203-213.

Dasnois, N. (2012) 'Uranium mining in Africa: A continent at the centre of a global nuclear Renaissance', *Governance of Africa's Resources Programme*, September 2012.

De Vita, G., Endersen, K., and Hunt, L. (2006) 'An empirical analysis of energy demand in Namibia', *Energy Policy,* 34(18), pp.3447-3463.

De Vita, G. and Trachanas, E. (2016) 'Nonlinear causality between crude oil price and exchange rate: A comparative study of China and India' - A failed replication (negative Type 1 and Type 2)', *Energy Economics, 56,* 150-160.

Deichmann, U., Meisner, C., Murray, S. and Wheeler, D. (2010) 'The economics of renewable energy expansion in rural Sub-Saharan Africa', *Policy Research Working Paper*, 5193.

Denisard, C., Rodrigo, D. L., and da Silveira, B. (2003) 'Short-run, long-run and cross elasticities of gasoline demand in Brazil', *Energy Economics*, pp.191-199.

Demierre, J., Bazilian, M., Carbajal, J., Sherpa, S. and Modi, V. (2015) 'Potential for regional use of East Africa's natural gas', *Applied Energy*, 143, pp. 414-436.

Diabi, A. (1998) 'The demand for electric energy in Saudi Arabia: an empirical investigation', *OPEC Review*, pp.13-24.

Dickey, D. and Fuller, W. A. (1979) 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association,* 74(366), pp.427- 431.

Dickey, D. and Fuller, W. A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root', *Econometrica,* 49, pp.1057-1072.

Dilaver, Z. and Hunt, L. C. (2011) 'Modelling and forecasting Turkish residential electricity demand', *Energy Policy*, pp.3117-3127.

Dong, Y. and Peng, C. J. (2013) 'Principled missing data methods for researchers', *SpringerPlus,* 2(222), pp.1-17.

Dramani, J. B. and Tewari, D. (2014) 'An econometric analysis of residential electricity demand in Ghana', *Mediterranean Journal of Social Sciences*, pp.209-222.

EIA (2007) *Annual Energy Outlook 2007*. Energy Information Administration, Washington, DC.

EIA (2013) *Oil and natural gas in Sub-Saharan Africa*.

EIA (2014) *South Africa Analysis,* pp. 1-13.

Eltony, M. (1996) 'An econometric study of the demand for gasoline in the Gulf Cooperation Council Countries', *The Journal of Energy and Development, 19*, pp.265-273.

Eltony, M. and Hoque, A. (1996) 'A cointegrating relationship in the demand for energy: The case of electricity in Kuwait', *Journal of Energy and Development, 21*, pp.293-301.

Endersen, K. (2004) *An empirical investigation of energy demand in Namibia*. PhD thesis. Oxford Brookes University.

Entorf, H. (1997) ‘Random walks with drifts: nonsense regression and spurious fixed-effect estimation’, *Journal of Econometrics,* 80, pp.287- 296.

Erdal, G., Erdal, H. and Esengun, K. (2008) ‘The causality between energy consumption and economic growth in Turkey’, *Energy Policy,* 36(10), pp.3838-3842.

Esso, L. J. and Keho, Y. (2016) ‘Energy consumption, economic growth and carbon emissions: Cointegration and causality evidence from selected African countries’*, Energy*, 114, pp. 492-497.

Expo, U., Chuku, C. A. and Effiong, E. L. (2011) ‘The dynamics of electricity demand and consumption in Nigeria: application of the bounds testing approach’, *Current REsearch Journal of Economic Theory*, pp.43-52.

Fishbone, L. and Abilock, H. (1981) ‘MARKAL, a linear-programming model for energy systems analysis: technical description of the BNL version’, *Energy Research, 5*, pp.353-375.

Fisher-Vanden, K., Jefferson, G. H., Liu, H. and Tao, Q. (2004) 'What is driving China's decline in energy intensity, *Resource and Energy Economics, 26*, pp.77-97.

Frimpong PB, Adu G. (2014); *Population Health and Economic Growth in Sub-Saharan Africa: A Panel Cointegration Analysis;* African Business, 15, pp.36-48.

Fuller, W. (1976) *Introduction to statistical time series.* New York: John Wiley and Sons.

Ghali, R. (1998) ‘The relationship between energy intensity and income levels: forecasting long term energy demand in Asian emerging countries’, *Energy 19*, pp.85-105.

Glaser, B. (1963) ‘Retrending research materials: The use of secondary analysis by the independent researcher’, *The American Behavioural Scientist,* 6(10), pp.11-14.

Glasure, Y. (2002) ‘Energy and national income in Korea: further evidence on the role of omitted variables’, *Energy Economics,* 24, pp.355-365.

Gnansounou, E., Bayem, H., Bednyagin, D and Dong, J. (2007) 'Strategies for regional integration of electricity supply in West Africa', *Energy Policy,* 35, pp.4142-4153.

Gould, J. P. and Lazear, E. P. (1989) *Microeconomic Theory.* United States of America: Richard D. Irwin.

Granger, C. and Newbold, P. (1974) 'Spurious regressions in econometrics' *Journal of Econometrics,* 2, pp.111-120.

Groen, J. and Kleibergen, F. (2003) 'Likelihood-based cointegration analysis in panels of vector error correction models', *Journal of Business and Economic Statistics,* 21, pp.295-318.

Gujarati, D. N. and Porter, D. C. (2009) *Basic Econometrics.* Fifth edn. Singapore: McGraw-Hill.

Harris, R. (1995) *Cointegration analysis in econometric modelling.* First edn. Great Britain: T.J Press (Padstow).

Hausman. (1978) 'Specification Tests in Econometrics' *Econometrica,* 46, pp.1251-1271.

Ho, C.Y. and Siu, K. W. (2007) 'A dynamic equilibrium of electricity consumption and GDP in Hong Kong: an empirical investigation', *Energy Policy,* 35(4), pp.2507-2513.

Hornby, W., Gammie, B. and Wall, S. (2001) *Business Economics.* Gosport: Pearson Education Limited.

Hsiao, C. (2014) *Analysis of Panel Data.* Third edn. United States of America: Cambridge University Press.

Hung, M. and Huang, T (2015) 'Dynamic demand for residential electricity in Taiwan under seasonality and increasing block pricing', *Energy Economics*, pp.168-177.

Hunt, C. (2011) 'Modelling and forecasting Turkish residential electricity demand', *Energy Policy*, pp.317-3127.

Holtedahl, P., and Joutz, F. L. (2004). 'Residential Electricity Demand in Taiwan', *Energy Economics*, 26, pp. 201-224.

Hwang, D. and Gum, B. (1991) 'The causal relationship between energy and GNP: the case of Taiwan', *Journal of Energy Development, 16*, pp.219-226.

IEA (2010) *World Energy Outlook 2010*.IEA, Paris.

IEA (2013) *World Energy Outlook.* Paris.

IEA (2014) *World Energy Outlook 2014 Special Report: Africa Energy Outlook*.

IEA (2015) Full report on China.

IEA (2015). *World Energy Statistics 2015 Revised Edition,* Paris.

International Energy Agency (2014): *Natural Gas Information : World Gas Supply 1960-2013* (Data downloaded: 25 October 2016). UK Data Service).

Im, K., Pesaran, M. H. and Shin, Y. (2003) 'Testing for unit roots in heterogeneous panels', *Econometrics,* 115, pp.53-74.

Inglesi- Lotz, R. (2011) 'The evolution of price elasticity of electricity demand in South Africa: a Kalman filter application', *Energy Policy*, pp.3690- 3696.

IRENA (2012) *Prospects for the African power sector.* Available at: http://www.irena.org/home.

Iwayemi, A. (2008) 'Nigeria's Dual Energy Problems: Policy Issues and Challenges', *International Association for Energy Economics.*

Iwayemi, A., Adenikinju, A. and Babatunde, A. A. (2010) 'Estimating petroluem products demand elasticities in Nigeria: A multivariate cointegration approach', *Energy Economics*, pp.73-85.

Iyke, B.N. (2015) 'Electricity consumption and economic growth in Nigeria: A revisit of the energy-growth debate', *Energy Economics*, 51, pp. 166-176.

Jamil, F. and Ahmad, E. (2010) 'The relationship between electricity consumption, electricity prices, and GDP in Pakistan' *Energy Policy*, pp.6016-6025.

Jebaraj, S. and Iniyan, S. (2006) ‘A review of energy models’, *Renewable and Sustainable Energy Reviews*, pp.281-311.

Jebli, M.B., Youssef, S.B. and Ozturk, I. (2015) ‘The Role of Renewable Energy Consumption and Trade: Environmental Kuznets Curve Analysis for Sub-Saharan Africa Countries’, *African Development Review*, 27, pp.288-300.

Johnson, M. (2014) ‘Secondary Data Analysis: A method of which the Time Has come’, *Qualitative and Quantitative Methods in LIbraries,* 3, pp.619-626.

Kalitsi, E. (2006) Hydropower development in Africa, Problems and Prospects. Available on http://www.un.org/esa/sustdev/sdissues/energy/op/nepadkalitsi_ppt.pdf (Accessed: 9 December, 2014).

Kan, H., Chen, B., Chen, C., Fu, Q. and Chen, M. (2004) ‘An evaluation of public health impact of ambient air pollution under various energy scenarios in Shanghai, China’, *Atmospheric Environment, 38*, pp.95-102.

Kanudia, A. and Loulou, R. (1999) ‘Advanced bottom -up modelling for national and regional energy planning in response to climate change’, *Environment and Pollution*, pp.191- 218.

Kao. (1999) ‘Spurious regression and residual-based tests for cointegration in panel data’, *Journal of Econometrics,* 90, pp.1-44.

Karekezi, S. (2002) ‘Poverty and energy in Africa- A brief review’, *Energy Policy,* 30, pp. 915-919.

Karimu, A. and Brannlund, R. (2013) ‘Functional form and aggregate demand elasticities: A nonparametric panel approach for 17 OECD countries’, *Energy Economics*, pp.19-27.

Karlsson, S. and Lothgren, M. (2000) ‘On the power and interpretation of panel unit roots’, *Economics Letters,* 66, pp.249-255.

Kaygusuz, K. (2012) ‘Energy for sustainable development: A case of developing countries’, *Renewable and Sustainable Energy Reviews,* pp. 1116-1126.

Kebede, E., kagochi, J. and Jolly, C. M. (2010) ‘Energy consumption and economic development in Sub- Sahara Africa’ *Energy Economics,* 32, pp.532-537.

Keho, Y. (2016) ‘What drives energy consumption in developing countries? The experience of selected African countries’, Energy Policy, 91, pp.233-246.

Kessides, I. (2014) ‘Powering Africa's sustainable development: The potential role of nuclear energy’, *Energy Policy*, 74, pp.557-570.

Khan, M. and Qayyum, A. (2008) ‘The demand for electricity in Pakistan’, *OPEC Energy Review*, pp.70-96.

Kialashaki, A. and Reisel, J. (2013) ‘Modelling of the energy demand of the residential sector in the United States using regression models and the artificial neural networks’, *Applied Energy*, pp.271-280.

Knight, K. and Schor, J. (2014) ‘Economic growth and climate change: A cross-national analysis of territorial and consumption-based carbon emissions in high income countries’, *Sustainability,* 6, pp.3722- 3731.

Kojima, M. and Matthews, W. (2010) *Petroleum Markets in Sub-Saharan Africa,* Washington. D.C: The World Bank.

KPMG (2013) Oil and gas in Africa, Africa’s *Reserves, Potential and Prospects.* KPMG Johannesburg.

Kraft, J. and Kraft, A. (1978) ‘On the relationship between energy and GNP’, *Journal of Energy and Development,* 3, pp.401- 403.

Kuma, S. (2008) ‘Cointegration and the demand for Energy in Fiji’, *Munich Personal RePEc Archive*, pp.1-16.

Lambe, F., Jürisoo, M., Wanjiru, H. and Senyagwa, J. (2015) ‘Bringing clean, safe, affordable cooking energy to households across Africa: an agenda for action’, *Stockholm Environment Institute, Stockholm and Nairobi, for the New Climate Economy*. Available at: http://newclimateeconomy.report/misc/working-papers.

Larsson, R., Lyhagen, J. and Lothgren, M. (2001) 'Likelihood-based cointegration tests in heterogeneous panels', *Econometrics,* 4, pp.109-142.

Lawler, K. and Seddighi, H. (1987) *Economic Theory and Modelling: An Integrated Approach.* Worcestershire: Peter Andrew Publishing Company.

Lee, C. and Chang, C. P. (2005) 'Structural breaks, energy consumption and economic growth revisited: evidence from Taiwan', *Energy Economics,* 27, pp.857-872.

Lescaroux, F. (2011) 'Dynamics of final sectoral energy demand and aggregate energy intensity', *Energy Policy*, pp.66- 82.

Levin, A., Lin, C and Chu, C. J. (2002) 'Unit root tests in panel data: asymptotic and finite-sample properties', *Econometrics,*108, pp.1-24.

Ley, K., Gaines, J. and Ghatikar, A. (2014) *The Nigerian Energy Sector -an Overview with a Special Emphasis on Renewable Energy, Energy Efficiency and Rural Electrification*. Germany: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

Loulou, R., Goldstein, G. and Noble , K. (2004) 'Documentation for the MARKAL family of models', *Energy Technology Systems Analysis Programme*, pp.1- 386. Paris.

Lundmark, R. (2001) 'Changes in Namibia's energy market', *Scandinavian Journal of Development*, pp.103- 112.

Maddala, G. S. and Wu, S. (1999) 'A comparative study of unit roots tests with panel data and a new simple test', *Oxford Bulletin of Economics and Statistics,* pp.631- 651.

Maladoh Bah, M and Azam, M. (2017) 'Investigating the relationship between electricity consumption and economic growth: Evidence from South Africa', *Renewable and Sustainable Energy Reviews*, 80, pp. 531-537.

Mallah, S. and Bansal, N. (2010) 'Allocation of energy resources for power generation in India: Business as usual and energy efficiency', *Energy Policy*, pp.1059-1066.

Mandelli, S., Barbieri, J. and Colombo, E. (2014) 'Sustainable energy in Africa: A comprehensive data and policies review', *Renewable and Sustainable Energy Reviews*, 37, pp.656-686.

Mangoyana, R. (2009) 'Bioenergy for sustainable development: An African context', *Physics and Chemistry of the Earth*, 34, pp.59- 64.

Marshall, A. (1980). *Principles of Economics: an introductory volume.* Cambridge: Macmillianand Co.

Masih, A. and Masih, R. (1996b) 'Energy consumption, real income and temporal causality: results from multi country based on cointegration and error correction modelling technique', *Energy Economics*, pp.165-183.

Masih, A. M. and Masih, R. (1997a) 'On the Temporal Causal Relationship between Energy Consumption, Real Income, and Prices: Some new Evidence from Asian-Energy Dependent NICs based on a Multivariate Cointegration/Vector Error-Correction Approach', *Policy Modeling,* 19(4), pp.417-440.

Masih, R. and Masih, A. M. (1996c) 'Stock-Watson dynamic OLS (DOLS) and error- correction modelling approaches to estimating long and short run elasticities in a demand function: New evidence and methodological implications from an application to the demand for coal in Mainland China', *Energy Economics*, pp.315-334.

McCoskey, S. and Kao, C. (1998) 'A residual-based test of the null of cointegration in panel data', *Econometric reviews,* 17, pp.57- 84.

Menegaki, A.N. and Tugcu. C.T. (2016) 'Rethinking the energy-growth nexus: Proposing an index of sustainable economic welfare for Sub-Saharan Africa', *Energy Research & Social Science*, 17, 147-159.

Mensah, J. T. (2014) 'Modelling demand for liquefied petroluem gas (LPG) in Ghana: current dynamics and forecast', *OPEC Energy Review*, 398-423.

Mensah, J.T., Marbuah, G. and Amoah, A. (2016) 'Energy demand in Ghana: A disaggregated analysis', *Renewable and Sustainable Energy Reviews*, pp.924-935.

Menyah, K. and Wolde-Rufael, Y. (2010) 'Energy consumption, pollutant emissions and economic growth in South Africa', *Energy Economics,* 32, pp.1374- 1382.

Mohammed, Y. and Eltony, M. (1996) 'The demand for natural gas in the Gulf Cooperation Council (GCC) states', *Middle East Business and Economic Review, 8*, pp.41- 48.

Morshed HAS. (2010); *A Panel Cointegration Analysis of the Euro area money demand;* Department of Statistics, Lund University.

Mulali, A. and Sab, C. N. (2012) 'The impact of energy consumption and CO2 emission on the economic growth and financial development in the Sub Saharan African countries', *Energy,* 39, pp.180- 186.

Murray, C. and Papell, D. H. (2000) 'Testing for unit roots in panels in the presence of structural change with an application to OECD unemployment', *Advances in Econometrics,* 15, pp.223-238.

Mukasa, A. D., Mutambatsere, E., Arvanitis, Y and Thouraya, T. (2013), Development of Wind Energy in Africa, Working Paper Series N° 170 African Development Bank, Tunis, Tunisia.

Mukasa, A. D., Mutambatsere, E., Arvanitis, Y and Thouraya, T. (2015) 'Wind energy in sub-Saharan Africa: Financial and political causes for the sector's under-development', *Energy Research and Social Science,* 5, pp. 90-104.

Narayan, P. and Narayan, S. (2010) 'Carbon dioxide emissions and economic growth: Panel data evidence from developing countries', *Energy Policy,* 38, pp.661- 666.

Narayan, P. and Smyth, R. (2005) 'The residential demand for electricity in Australia: an application of the bounds testing approach to co-integration', *Energy Policy*, pp.467-474.

Nelson, C. and Plosser, C. R. (1982) 'Trends and random walks in macroeconomic time series', *Journal of Monetary Economics,*10, pp.139-162.

Newiak, M. (2016) ‘Weakening Growth in Sub-Saharan Africa Calls for Policy Reset.’ *IMFSurvey Magazine: Countries and Regions*, 3 May.

Nigeria Central Bank, CBN. (2014) *Statistical Bulletin,* Abuja: CBN Statistical Office.

Odhiambo, N, O. (2009) ‘Electricity consumption and economic growth in South Africa: A trivariate causality test’, *Energy Economics,* 31, pp.635- 640.

Oh, W. and Lee, K. (2004) ‘Causal relationship between energy consumption and GDP: the case of Korea 1970- 1999’, *Energy Economics,* 1(26), pp.51- 59.

Oil and Gas Journal. (2013) ‘Worldwide Look at Reserves and Production, OandGJ’, Pennwell Corporation, Oklahoma City, Oklahoma.

Okafor, E. (2008) ‘Development crisis of power supply and implications for industrial sector in Nigeria’, *Studies of Tribes and Tribals,* 6 (2), pp.83-92.

Onyeji, I. (2014) ‘Harnessing and integrating Africa’s renewable energy resources’, *New Energy Insights*, 3, pp.27-38.

Onyeji, I. and Jones, E. L. (2014) ‘Harnessing and Integrating Africa's Renewable Energy Resources’, *Renewable Energy Integration*, pp 27- 36.

Ouedraogo, I, M. (2010) ‘Electricity consumption and economic growth in Burkina Faso: A cointegation analysis’, *Energy Economics,* 32, pp.524-531.

Ouedraogo, N. (2013) ‘Energy consumption and economic growth: Evidence from the economic community of West African States (ECOWAS)’, *Energy Economics,* 36, pp.637- 647.

Ouedraogo, N. (2017) ‘Africa energy future: Alternative scenarios and their implications for sustainable development strategies’, Energy Policy, 106, pp. 457-471.

Owen, M., Van der Plas, R. and Sepp, S. (2013) ‘Can there be energy policy in Sub-Saharan Africa without biomass?’, *Energy for Sustainable Development,* 2, pp.146-152.

Oyedepo, S. (2012) ‘On energy for sustainable development in Nigeria’, *Renewable and Sustainable Energy*, pp.2583- 2598.

Ozturk, I. (2010) ‘A literature survey on energy-growth nexus’*, Energy Policy,* 38, pp.340 -349.

Ozturk, I., Aslan, A. and Kalyoncu, H. (2010) ‘Energy consumption and economic growth relationship: Evidence from panel data for low and middle income countries’, *Energy Policy,* 38, pp.4422- 4428.

Pandey, R. (2002) ‘Energy policy modelling: agenda for developing countries’, *Energy Policy*, pp.97-106.

Paul, S. and Bhattacharya, R. N. (2004) ‘Causality between energy consumption and economic growth in India: a note on conflicting results’, *Energy Economics,* 26(6), pp. 977- 983.

Payne, J. (2009) ‘On the dynamics of energy consumption and output in the US’, *Applied Energy,* 86(4), pp.575- 577.

Pedroni, P. (2000) ‘Fully modified OLS for heterogeneous cointegrated panels’, *Advances in Econometrics,* 15, pp. 93-130.

Pedroni. (2004) ‘Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis’, *Econometric Theory,* 20, pp. 597- 625.

Pesaran, M. (2004) ‘General diagnostic tests for cross-section dependence in panels, *working paper.*

Pesaran, M., Akiyama, T, and Smith, R. (1998) ‘*Energy Demand in Asian Developing Economies.* Oxford: Oxford University Press.

Pesaran, M., Shin, Y. and Smith, R. (1999) ‘Pooled mean group estimation of dynamic heterogeneous panels’, *Journal of American Statistical Association,* 94, pp. 621- 634.

Pesaran, M., Smith, R. and Akiyama, T. (1998). *Energy demand in Asian economies.* Oxford: Oxford University Press.

Persyn, D. and Westerlund, J. (2008) 'Error-correction–based cointegration tests for panel data', *Stata Journal,* 2, pp. 232-241.

Phillips, P. and Moon, H. (1999) 'Linear regression limit theory for nonstationary panel data', *Econometrica,* 67, pp.1057-1111.

Phillips, P. and Perron, P. (1988) 'Testing for a unit root in time series regression',*Biometrika,* 75(2), pp. 335-346.

Poumanyvong, P. and Kaneko, S. (2010) 'Does Urbanisation lead to less energy use and lower CO2 emissions? A cross-country analysis.', *Ecological Economics,* 70, pp.434- 444.

PRB (2013) *World Population Data sheet.* Available at: http://www.prb.org/publications/datasheets/2013/2013-world-population-data-sheet/data-sheet.aspx [Accessed 15 April 2016].

Rahut, D. B., Behera, B. and Ali, A. (2016) 'Patterns and determinants of household use of fuels for cooking: Empirical evidence from sub-Saharan Africa', *Energy*, 117, pp.93-104.

Rahut, D. B., Behera, B. and Ali, A. (2017) 'Factors determining household use of clean and renewable energy sources for lighting in Sub-Saharan Africa', *Renewable and Sustainable Energy Reviews*, 72, pp. 661-672.

Rubin, D. (1976) 'Inference and Missing Data', *Biometrika,* 63, pp. 581- 592.

Rysankova, D., Putti, V.R., Hyseni, B., kammila, S. and Kappen, J.F. (2014) Clean and improved cooking in Sub-Saharan Africa: A landscape Report. Report No. 98664. Africa Clean Cooking Initiative. The World Bank, Washington DC. Available at: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2015/08/18/090224b08307b414/4_0/Rendered/PDF/Clean0and0impr000a0landscape0report.pdf

Sadorsky, P. (2013) 'Do urbanisation and industrialization affect energy intensity in developing countries?', *Energy Economics,* 37, pp. 52- 59.

Saed, J. (2004) 'Estimation of total energy consumption in Jordan 1980-1999: Evidence from cointegration and Stock', *Journal of Economic and Administrative Sciences*, pp.71-82.

Samuel, Y., Manu, O. and Wereko, T. (2013) 'Determinants of energy consumption: A review', *International Journal of Management Sciences*, pp.482-487.

Samuelson, P. and Nordhaus, W. (1992) *Micro-Economics.* New York: McGRAW-Hill, INC.

Sander, K., Halder, S.W and Hyseni, B. (2011) *Wood-Based Biomass Energy Development for Sub-Saharan Africa: Issues and Approaches*, Washington, DC: The International Bank for Reconstruction: World Bank Group.

SAPP (2013) *Southern African Power Pool yearly Statistics*, SAPP, Harare, Zimbabwe.

SAPP (2016) *Southern African Power Pool yearly report*, SAPP, Harare, Zimbabwe.http://www.sapp.co.zw/docs/SAPP%20ANNUAL%20REPORT%202016.pdf//. Accessed on 1 March, 2017.

Schafer, J. (1999) 'Multiple imputation: a primer', *Statistical methods in Medical research,* 8(1), pp.3-15.

Schafer, J. L. and Graham, J. W. (2002) 'Missing data: Our view of the state of the art', *Psychological methods,* 7(2), pp.147-177.

Shahbaz, M., Solarin, S.A., Sbia, R. and Bidi, S. (2015) 'Does energy intensity contribute to CO2 emissions? A trivariate analysis in selected African countries', *Ecological Indicators*, 50, pp.215- 224.

Sharma, D., Nair, C. and Balasubramanian. (2002) 'Demand for commercial energy in the state of Kerala, India: an econometric analysis with medium-range projections', *Energy Policy*, pp.781- 791.

Shrestha, R. and Rajbhandari, S. (2010) 'Energy and environmental implications of carbon emission reduction targets: case of Kathmandu Valley, Nepal', *Energy Policy, 38*, pp.4818-4427.

Shukla, P. (1995) ‘Greenhouse gas models and abatement costs for developing nations: A critical assessment’, *Energy Policy*, pp.677-687.

Sinton, J. and Fridley, D. (2000) ‘What goes up: recent trends in China's energy consumption’, *Energy Policy, 28*, pp.671- 687.

Skea, J. (2014) Findings *of the IPCC Fifth Assessment: what they mean for energy*: BIEE Seminar.

Smith, E. (2008) *Using secondary data in educational and social research.* New York: McGraw-Hill education.

Soytas, U. and Sari, R. (2009) ‘Energy consumption, economic growth, and carbon emissions: challenges faced by an EU candidate member’, *Ecological Economics,* 68(6), pp. 1667- 1675.

Suganthi, L. and Samuel, A. (2012) ‘Energy models for demand forecasting- A review’, *Renewable and Sustainable Energy Reviews*, pp.1223-1240.

Sulaiman, C., Abdul-Rahim, A.S., Chin, L. and Mohd-Shahwahid, H.O. (2017) ‘Wood fuel consumption and mortality rates in Sub-Saharan Africa: Evidence from a dynamic panel study’ *Chemosphere,* 177, pp. 224-231.

Sun, Y. (2004) ‘Estimation of the long-run average relationship in nonstationary panel time series’, *Econometric theory,* 20, pp.1227-1260.

Swan, L. and Ugursal, V. I. (2009) ‘Modeling of end-use energy consumption in the residential sector: A review of modeling techniques’, *Renewable and Sustainable Energy Reviews*, pp.1819-1835.

Tabachnick, B.G and Fidell, L.S. (2007), *Using Multivariate Statistics*, 5th edn. Boston: Pearson Education, Inc.

Tabachnick, B. and Fidell, L. S. (2012) *Using multivariate statistics.* 6th ed. Needham Heights, MA: Allyn and Bacon.

Tekaya, A. (2008) *The impact of asset specificity on outsourcing performance: A dissaggregated analysis by buyer-supplier asset specificity dimensions*. PhD thesis. Oxford Brookes University.

The Economist (2016) Manufacturing: Not making it. *The Economist*, 16th-22nd April, pp.3-18.

The Windpower, Wind turbines and wind farm database, http://www.thewindpower.net/country_africa_en.php, accessed on 8 December, 2014.

The World Bank. Global gas flaring reduction partnership, http://go.worldbank.org/NEBP6PEHS0, accessed on 11 December, 2014.

Torriti, J. (2014) 'A review of time use models of residential electricity demand', *Renewable and Sustainable Energy Reviews*, pp.265-272.

Tsai, M. and Chang, S. (2013) 'Taiwan's GHG mitigation potentials and costs: An evaluation with the MARKAL model', *Renewable and Sustainable Energy Reviews*, pp.294-305.

Tunc, M., Camdali, U. and Parmaksizoglu, C. (2006) 'Comparison of Turkey's electrical energy consumption and production with some European countries and optimization of future electrical power supply investments in Turkey', *Energy Policy, 34*, pp.50-59.

Uddin, N. and Taplin, R. (2015) Regional cooperation in widening energy access and also mitigating climate change: current programs and future potential *Global Environmental Change (Downloaded: July 2015).*

UKData Service (2015) *International Macrodata*. Available at: https://www.ukdataservice.ac.uk/get-data/key-data/international-macro-databanks [Accessed 15 October 2015].

UNFPA (2012) Status Report, Adolescents and Young People in Sub-Saharan Africa; Opportunities and Challenges.

UNECA (2006) Energy for sustainable development in Africa. CSD14. UN Economic commission for Africa.

UNICEF (2014) In: Division of data, research and policy report on Generation 2030 Africa, United Nations International Children Emergency Fund.

Urban, F., Benders, R. M. and Moll, H. C. (2007) 'Modelling energy systems for developing countries', *Energy Policy*, pp.3473-3482.

USGS (United States Geological Survey) (2000) *World Petroleum Assessment'*, USGS Boulder, United States.

USGS (United States Geological Survey), (2012a). *'An estimate of undiscovered oil and Gas resources of the world'*, fact sheet FS2012-3042, USGS, Boulder, United States.

USGS (United States Geological Survey), (2012b) *Assessment of potential additions to conventional oil and gas resources of the world (outside the United States) from reserves growth* fact sheet 2012-3052, USGS, Boulder, United States.

Verbeek, M. (2012) *A guide to modern econometrics.* 4th ed. West Sussex: Wiley and Sons Ltd.

Victor, D. (2009) The politics of fossil-fuel subsidies, Global Subsidies Initiative (GSI) of the International Institute for Sustainable Development (IISD). IISD, Geneva

Vines, A. (2012) 'Opportunities and challenges in Africa's changing energy landscape', *Oxford Energy Forum,* 90, pp.10-12.

Wagner, A. (2014) *International Fuel Prices 2012/2013,* Berlin: Deutsche Gesellschaft fur Internationale Zusammenarbeit (GIZ) GmbH.

Wagpco (West Africa Gas Pipeline Company), 'The pipeline system', Avialable at: http://www.wagpco.com/index.php?option=com_contentandview=articleandid=114andItemid=105andlang=en (Accessed: 8 December 2014)

Wang, Z., Zeng, H., Wei, Y. and Zhang, Y. (2012) 'Regional total factor energy efficiency: An empirical analysis of industrial sector', *Applied Energy,* 97, pp.115- 123.

Wesseh, P.K. and Lin, J.P. (2017) ‘Options for mitigating the adverse effects of fossil fuel subsidies removal in Ghana’, *Journal of Cleaner Production,* 141, pp.1445-1453.

Westerlund, J. (2007) ‘Testing for error correction in panel data’, *Oxford Bulletin of Economics and Statistics,* 69, pp.709- 748.

Wicke, B., Smeets, E., Watson, H. and Faaij, A. (2011) ‘The current bioenergy production potential of semi-arid and arid regions in sub-Saharan Africa’, Biomass *and Energy*, 35, pp. 2773- 2786.

Wolde- Rufael, Y. (2004) ‘Disaggregated industrial energy consumption and GDP: the case of Shanghai’, *Energy Economics,* 26, pp. 69-75.

Wolde-Rufael (2009) ‘Energy consumption and economic growth: The experience of African countries revisited’, *Energy Economics,* pp.217-224.

Wolde-Rufael, Y. (2006) ‘Electricity consumption and economic growth: a time series experience for 17 African countries’, *Energy Policy,* pp.1106-1114.

Wolde-Rufael. (2006) ‘Disaggregated industrial energy consumption and GDP: the case of Shanghai 1952- 1999’, *Energy Economics*, pp.69-75.

Wolfram, C., Shelef, O. and Gertler, P. J. (2012) ‘How will energy demand develop in the developing world?’, *NBER working papers series,* pp.1- 21.

World Bank(2015) *Data overview.*

World Bank (2016) ‘*Global Outlook: Divergences and risks- June 2016’.* Available at: http://pubdocs.worldbank.org/en/154911463605617095/Global-Economic-Prospects-June-2016-Global-Outlook.pdf. (Accessed: 25 October 2016).

World Development Indicators, World Bank. Available at: http://data.worldbank.org/region/SSA (Accessed: 5 December 2014).

World Nuclear Association (2014) *Uranium in Namibia*. Available at: http://www.world-nuclear.org/info/Country-Profiles/Countries-G-N/Namibia/ (Accessed: 9 December 2014).

World Nuclear Association (2014) *Uranium in Niger*. Available at: http://www.world-nuclear.org/info/country-profiles/countries-g-n/niger/ (Accessed: 9 December 2014).

World Nuclear Association (2014) *Uranium in South Africa*. Available at: http://www.world-nuclear.org/info/Country-Profiles/Countries-O-S/South-Africa/ (Accessed: 9 December 2014).

World Nuclear association (2016) *Uranium in Africa*. Available at: http://www.world-nuclear.org/information-library/country-profiles/others/uranium-in-africa.aspx (Assessed: 26 October 2016).

Wu, Y. (2012) 'Energy intensity and its determinants in China's regional economies', *Energy Policy,* 41, pp.703- 711.

Young, R. and Johnson, D. R. (2015) 'Handling missing values in longitudinal panel data with multiple imputation', *Journal of Marraige and Family,* 77, pp.277-294.

Yu, E. and Jin, J. C. (1992) 'Cointegration tests of energy consumption, income and employment', *Resources and Energy,* 14, pp.259- 266.

Yule. G. (1926) 'Why do we Sometimes get Nonsense-Correlations between Time-Series?-A Study in Sampling and the Nature of Time-Series', *Journal of the Royal Statistical Society,* 89(1), pp.1-63.

Zachariadis, T. and Pashourtidou, N. (2007) 'An empirical analysis of electricity consumption in Cyprus', *Energy Economics*, pp.183-198.

Zarnikau, J. (2003) 'Functional forms in energy demand modelling', *Energy Economics, 25*, pp.603- 613.

Zeng, L, Xu, M, Liang, S, Zeng, S and Zhang, T. (2014) 'Revisiting drivers of energy intensity in China during 1997- 2007: A structural decomposition analysis', *Energy Policy,* 67, pp.640- 647.

Zhang, X. and Cheng, X. (2009) 'Energy consumption, carbon emissions and economic growth', *Ecological Economics,* 68(10), pp.2706-2712.

Ziramba, E. (2008) 'The demand for residential electricity in South Africa', *Energy Policy, 36*, pp.3460-3466.

Appendix 1

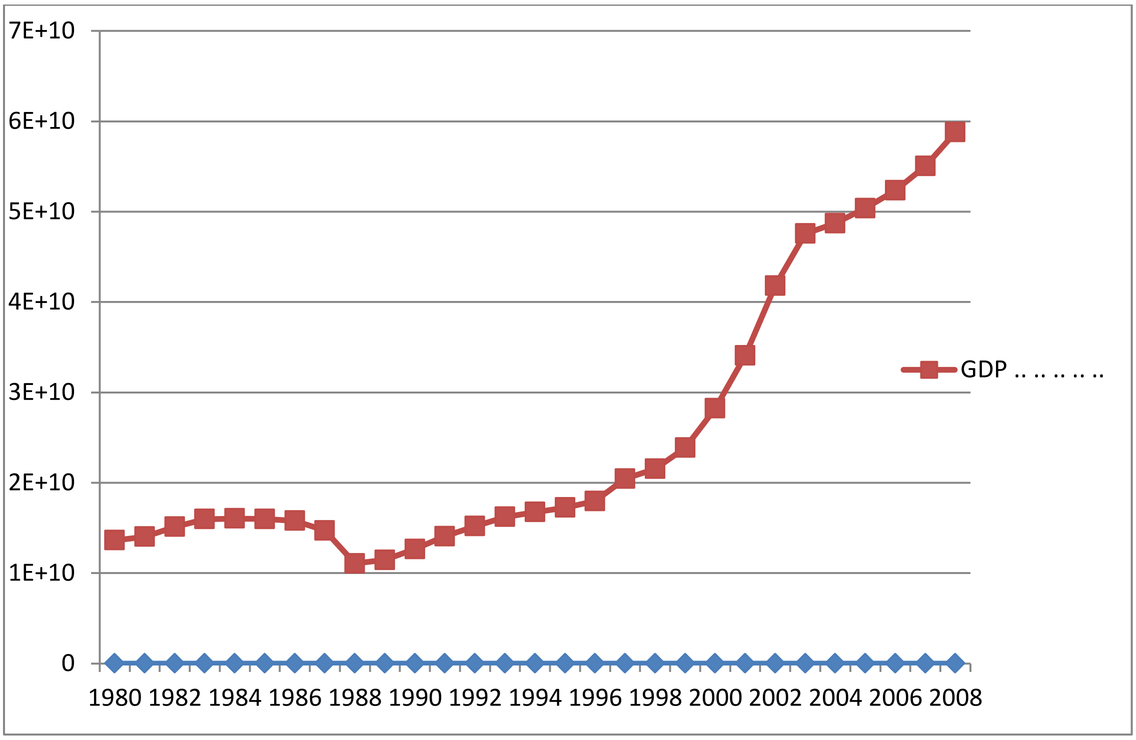

Figure a: Angola GDP between 1980-2013

(**Source:** created by author using SSA dataset)

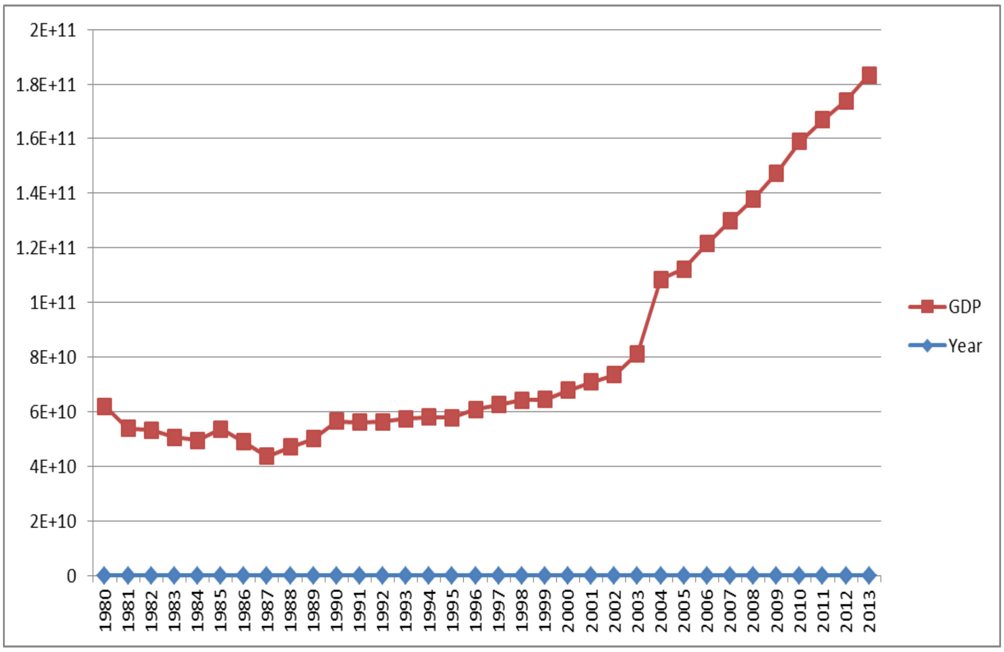

Figure b: Nigeria GDP between 1980-2013

(**Source:** created by author using SSA dataset)

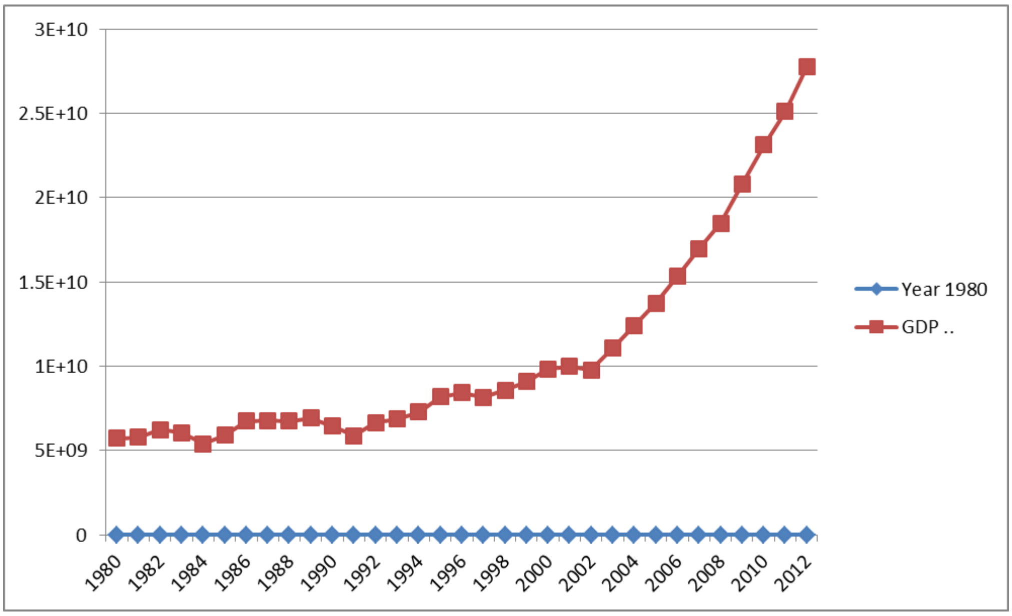

Figure c: Ethiopia GDP between 1980-2013

(**Source:** created by author using SSA dataset)

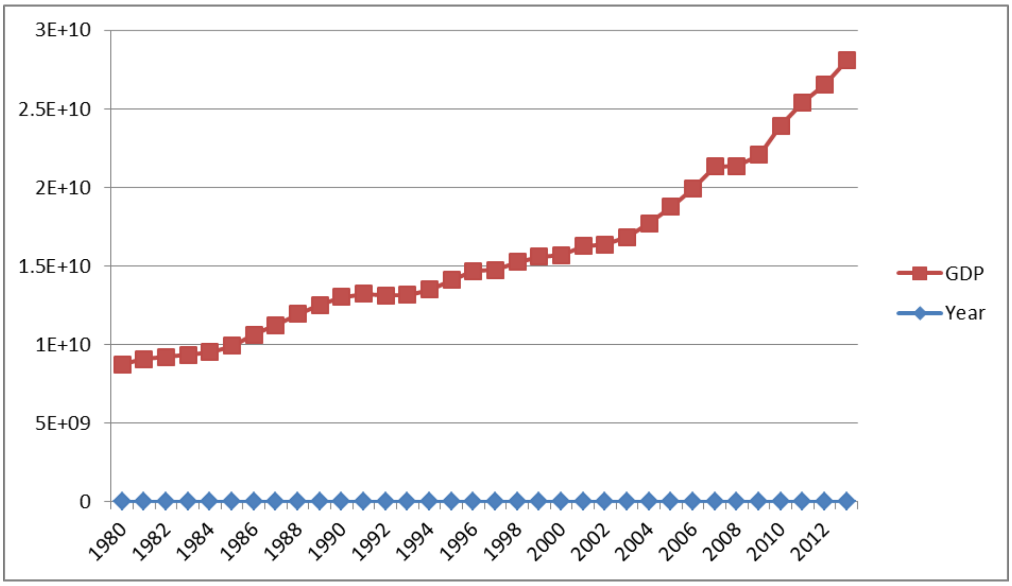

Figure d: Kenya GDP between 1980-2013

(**Source:** created by author using SSA dataset)

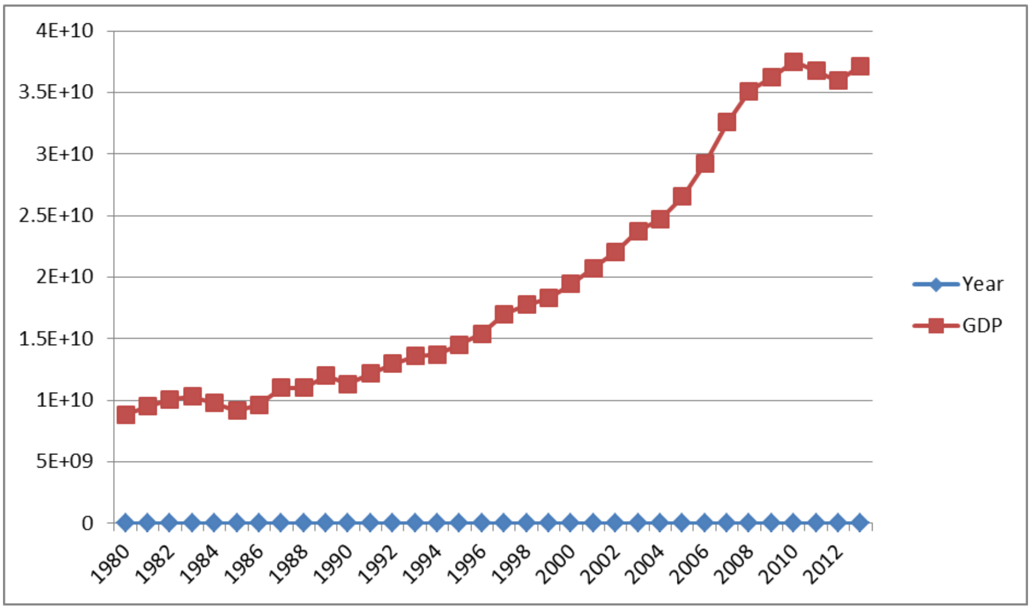

Figure e: Sudan GDP between 1980-2013

(**Source:** created by author using SSA dataset)

Appendix 2

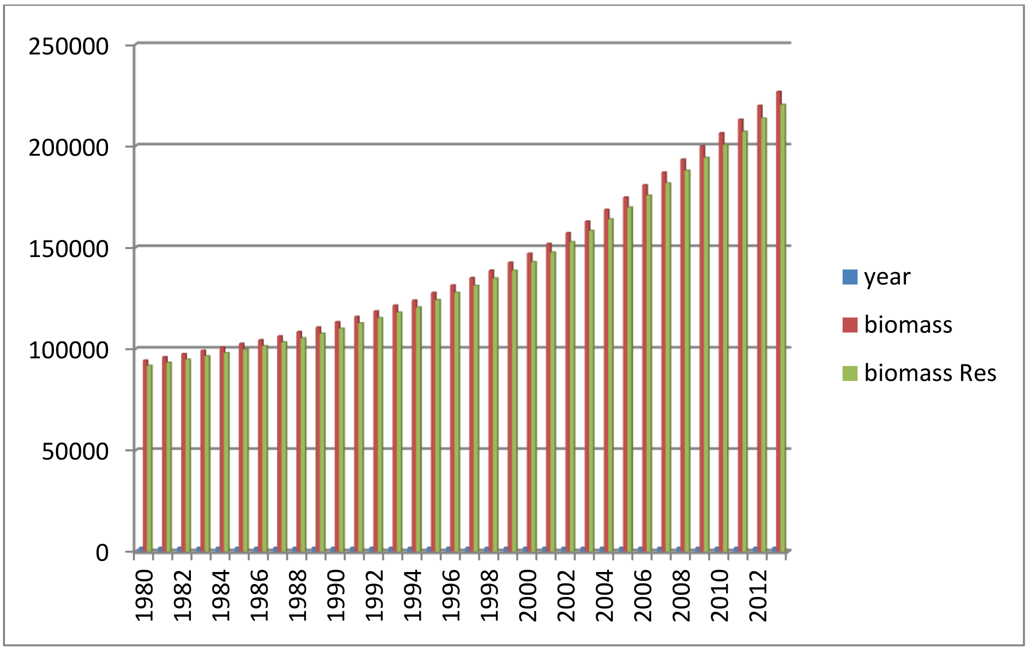

Figure a: Angola biomass consumption

(**Source:** created by author using SSA dataset)

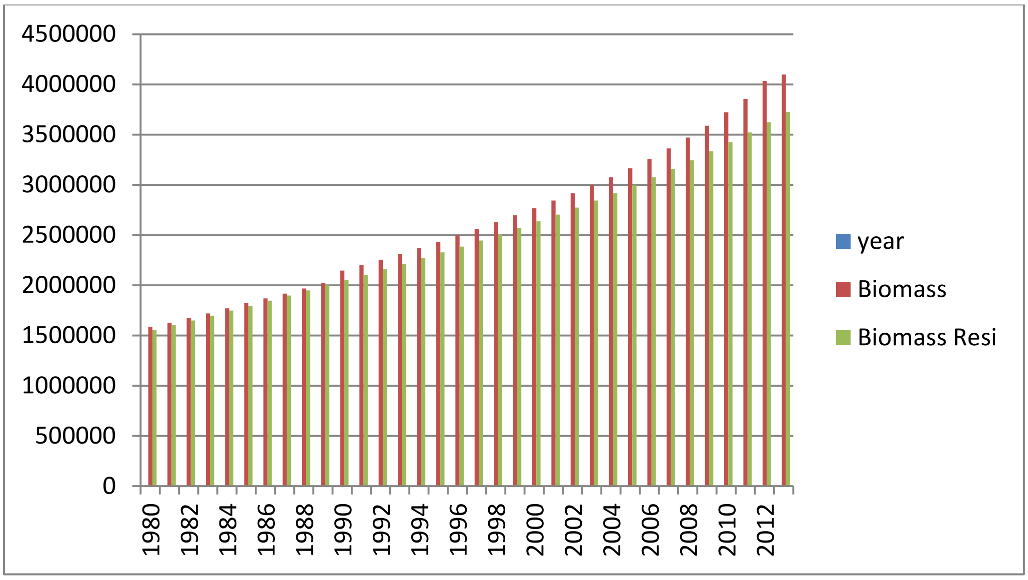

Figure b: Nigeria biomass consumption

(**Source:** created by author using SSA dataset)

Note: biomass represents the total biomass consumption; biomass resi stands for the total biomass consumption in the residential sector by consumers.