Effects of Malaria on Farmers' Technical Efficiency In Africa

Simon Oluwaseun Olayiwola Soname

Thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

School of Agriculture, Policy, and, Development

Department of Agricultural and Food Economics

May, 2016

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Simon .O.O. Soname

May, 2016

Abstract

Malaria is a problem in Africa. Thus, the aim of this research is to present a *reliable* measure of the farmers' *Willingness-To-Pay* for malaria abatement in Africa.

We develop a model that inputs the stochastic frontier model into the household production model. We analyse our model using the Bayesian Markov chain Monte Carlo techniques of Gibbs Sampling and Metropolis-Hastings. In order to arrive at a reliable posterior distribution for our Willingness To Pay, we multiply the individual households' price for agricultural staples by the corresponding malaria estimates from our analysis together with a constant value.

We apply our model to datasets from Nigeria, Ethiopia, and, Tanzania and report the corresponding Willingness To Pay point estimates and posterior distributions. Our results show that on the average, farmers in these three countries are willing to pay less than US$1 for a 100 per cent increase in malaria case per 1000 individual per annum.

Policy makers can use these values to introduce minimum prices and gradual repayment schemes for prophylactic measures.

ACKNOWLEDGEMENTS

Contents

List of Tables

List of Figures

1. Overview

In this chapter, we present the overview of our research; thus, this chapter is central to this thesis, at the end of which, we expect the reader to fully understand the reason why this research is relevant at this time and its contribution to knowledge in this area.

The household is the basic unit of the society and production, and, one of the most important resources of the household as emphasized by researchers - Mincer et al. 1963; Becker 1965; Singh et al. 1986 - is the total available time which they have to rationally share between productive activities and leisure. However, one resource, a possible competitor on the scale of importance to the household, is the *health* of its members. Good health is a minimum requirement for successful labour intensive farming, hence, preventing losses due to poor health would be valuable to the household.

In this research, we take health as an input which competes with other inputs on the degree of significance to the household. We focus on prevention of malaria infection as a health input, which the household has to wish for and purchase prophylactic drugs for. Malaria is a major health issue facing rural farmers in Africa. It is a disease that has been with man from time immemorial and Africa has the largest prevalence rate in the world with almost 90% of cases occurring there (Breman et al. 2004) (see also figure 1.1, 1.2, and, 1.3 at the end of this chapter). The ability of the household to take decisions on a particular malaria control measure might be the key to sustainable food security and production. This serves as a great motivation to investigate how much the household would be willing to pay for services that militate against malaria infection. We take the malaria parasite prevalence rate as a motivation for the willingness of the household to pay or not to pay for prophylactic measures.

Malaria is a vector borne disease caused by the parasite Plasmodium and its vector is the mosquito. There are different species of Plasmodium but the common ones in Africa are the *Plasmodium falciparum* and *Plasmodium vivax*. We shall dwell more on these two species of malaria parasite later in chapter two. Just like the parasites, its vector, mosquito, also has different species and strains, but the most common ones are *Anopheles gambiae* group and *Anopheles funestus* group.

The main aim of this research is to estimate the effect of malaria on rural farmers' technical efficiency in Africa. In achieving this objective, we seek to precisely estimate the willingness of rural farming households in Africa to pay for prophylactic measures like the insecticide treated nets, use of prophylactic drugs, and, indoor residual spraying. The

way and manner this is done forms the fulcrum of this research. Prevention of malaria is likely to increase the amount of time available for work, increase production, farm income and using the words of Holloway and Ehui (2001) will "increase market penetration and enhance the surplus generating potential of the household" especially in a continent where attaining food availability and security is still a conundrum.

In placing a value on the amount the household will be willing to pay for malaria parasite mitigation we use rich data sets from the World Bank Living Standard Survey for Nigeria, Tanzania and the International Food Policy and Research Institute (IFPRI) Ethiopian Rural Household Survey. One great characteristic of these data sets is the collection of the Geo-location variables of the households as the survey was carried out which enables us to over-lay the spatial malaria data from the Malaria Atlas Project (M.A.P) over them. Also, each of the household data contains similar quantities of interest that enable us to compare results on a country basis. The innovative art of merging these two separate data sets, we believe, is one of the strengths of this research.

We experience certain difficulties in the use of these data sets and primary among them is the merger of both the malaria data and the household survey data. This is because the data sets come in different electronic configurations and thus, requires the application of certain computational and software skills which we will discuss later in this thesis. Another constraint is the fact that the years in which the data were collected are different and also there are variations in the way each of the data were collected. We could not do anything about this apart from analysing the data the way they are.

Another notable aspect of this research is the linkage between the theoretical framework and the analytical framework. We start our research with a modification of the basic household model (Singh et al. 1986); develop a composed error model around it by robustly exploring and utilizing the known conditional distributions of the important unknown quantities and substituting the quantities derived from this model into our Willingness-To-Pay model in order to obtain a value for the quantity which we desire. The composed-error model enables us to obtain productivity and efficiency values which can be used to compare efficiency across different zones, regions, countries and across time dimensions. We then substitute the inefficiency values for malaria prevalence in our household model, in short, our inability to obtain our inefficiency values will preclude our calculation of the willingness – to – pay estimate.

The use of the Bayesian technique and principles in our analysis helps to mitigate a lot of the issues raised above. The technique also makes the utilization and exploration of our data sets easier. But this comes at a price, this is because the use of the Bayesian techniques requires a lot of programming, simulation, and computing time. The simulation technique we employ is the Markov Chain Monte Carlo techniques of Gibbs sampling and the Metropolis - Hastings algorithm. To properly understand this technique and in order to arrive at unbiased estimates, we undertake exercises in MATLAB$^{\copyright}$, the codes of which

2

we present in the appendix to this thesis. Part of the exercises involve the calculation of the marginal likelihood. This quantity confers an opportunity to compare models and also to select the best model from a group of models. A model selection code, which is not part of the exercises (we thank Dr. Garth Holloway for providing us with this code and instigating the Compensating Variation framework), enables us to test if the malaria prevalence variable will be among the prime variables needed for our composed error analysis.

We run the model selection on the different forms of the normal linear production models which we present in the appendix. The model selection code, however, did not select malaria prevalence as a prime variable in all our analyses; this could be due to data problems, which we will discuss in later chapters. But the prime objective of this research still remains, which is, given a well-behaved household data set and given that malaria affects farmers' technical efficiency, we could exactly estimate the amount a particular household is willing to pay for malaria mitigation. Arriving at a reliable quantity will give policy makers the opportunity to exactly know how much *subsidy* should be given to the *household* (if there will be), in order, for them to patronise a particular prophylactic measure, for example, insecticide treated nets. It is also expected that if households pay for a particular prophylactic measure, there is the tendency for them to see to the success of that particular preventive programme as they are likely to see it more like their own.

The rigorous use of the Bayesian methodology furnishes us with an acceptable prediction on how much the household is willing to pay in the future if malaria prevalence changes (or remains constant) by certain values in the future. In other words, we are able to state how much a non-existing household will be willing to pay for malaria mitigation in the future.

All of the above forms the crux of the remaining chapters of this research. The reader should, however note that what we seek to quantify is the direct cost of malaria to the household and not the sum of the direct, indirect and intangible costs of malaria to the household.

The rest of this thesis is divided into nine other chapters. Presently, in chapter two, we present the introduction to this study, our problem statement, specific objectives, and, malaria epidemiology. In chapter three, we present empirical evidence on the household model, chapter four explains our theoretical framework, while, we peruse the literature on the composed-error model in chapter five.

In chapter six, we espouse empirical evidence on the Markov chain Monte Carlo diagnostics; chapter seven discusses the data. In chapter eight, we show our empirical framework, while in chapter nine, we present our results and discuss them. In chapter ten, we conclude and make recommendations on this research.

Mean point estimates of the age-standardised annual mean *Plasmodium falciparum* parasite rate in two to ten year olds (*Pf*PR$_{2\text{-}10}$) within the spatial limits of stable transmission. Areas of no risk and unstable risk (*Pf*API < 0.1‰) are also shown.
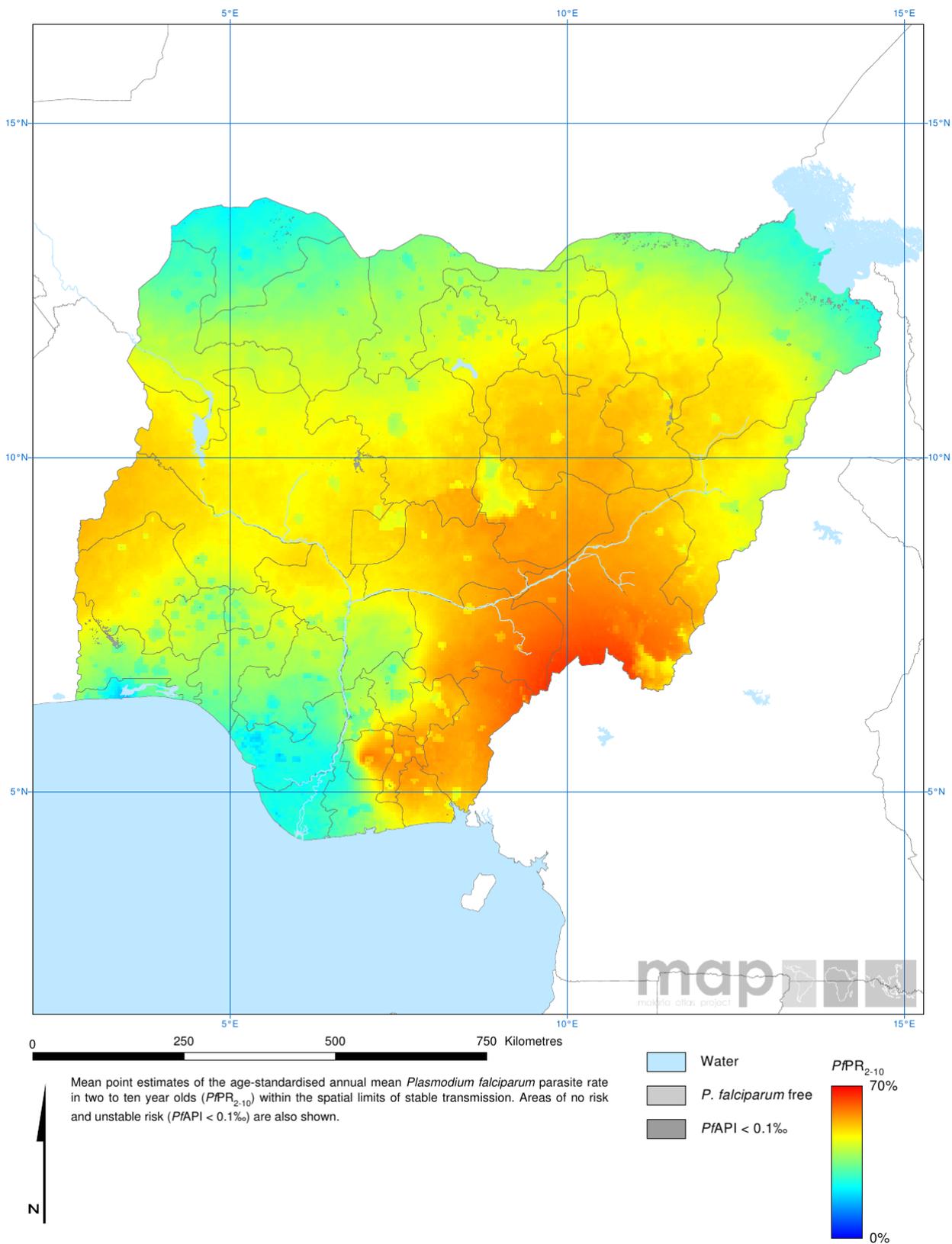
Figure 1.1.:   Spatial Distribution of Malaria Endemicity In Nigeria, 2010

Adapted from Gething et al. (2011)
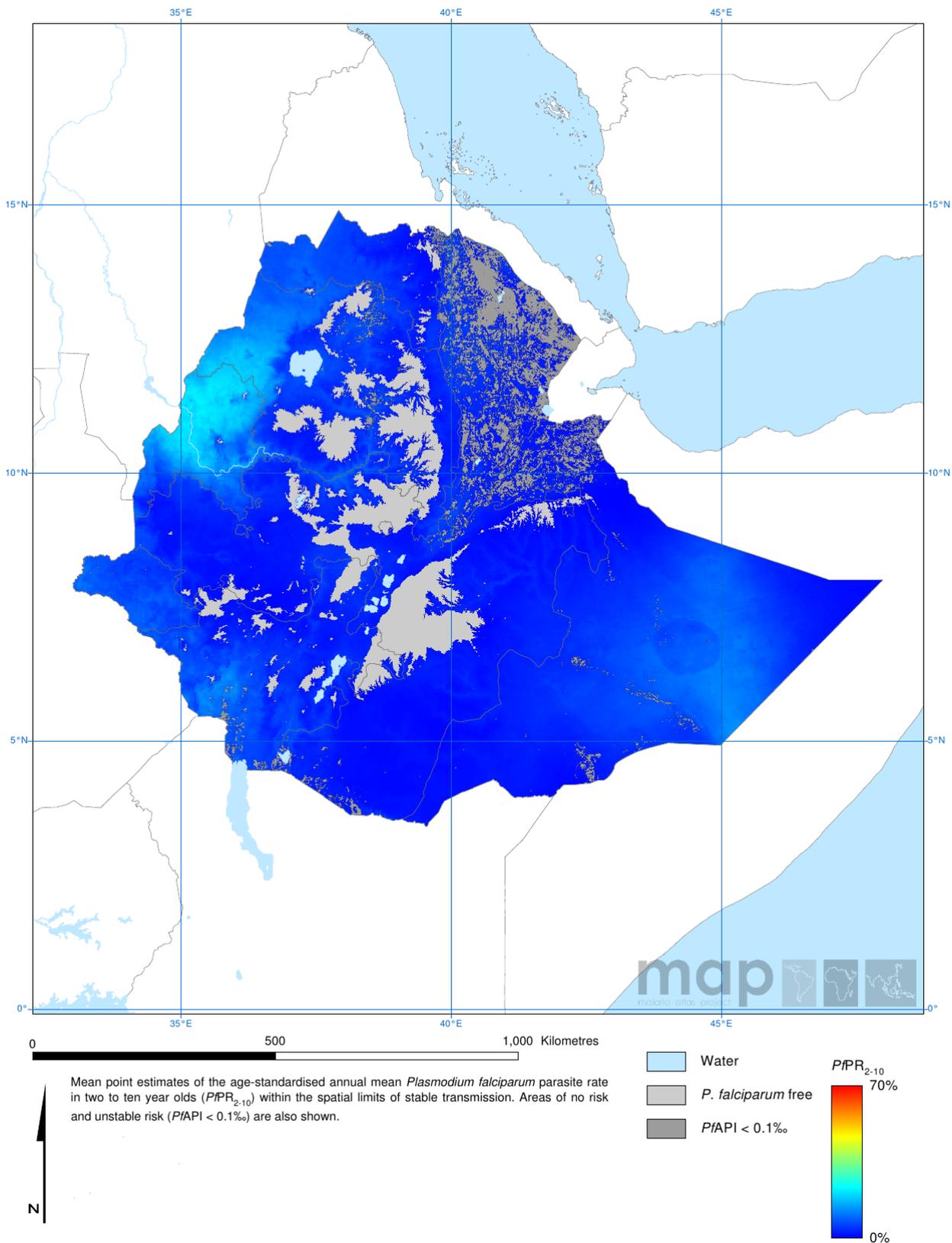
Figure 1.2.:   Spatial Distribution of Malaria Endemicity In Ethiopia, 2010

Adapted from Gething et al. (2011)

Mean point estimates of the age-standardised annual mean *Plasmodium falciparum* parasite rate in two to ten year olds (*Pf*PR$_{2\text{-}10}$) within the spatial limits of stable transmission. Areas of no risk and unstable risk (*Pf*API < 0.1‰) are also shown.
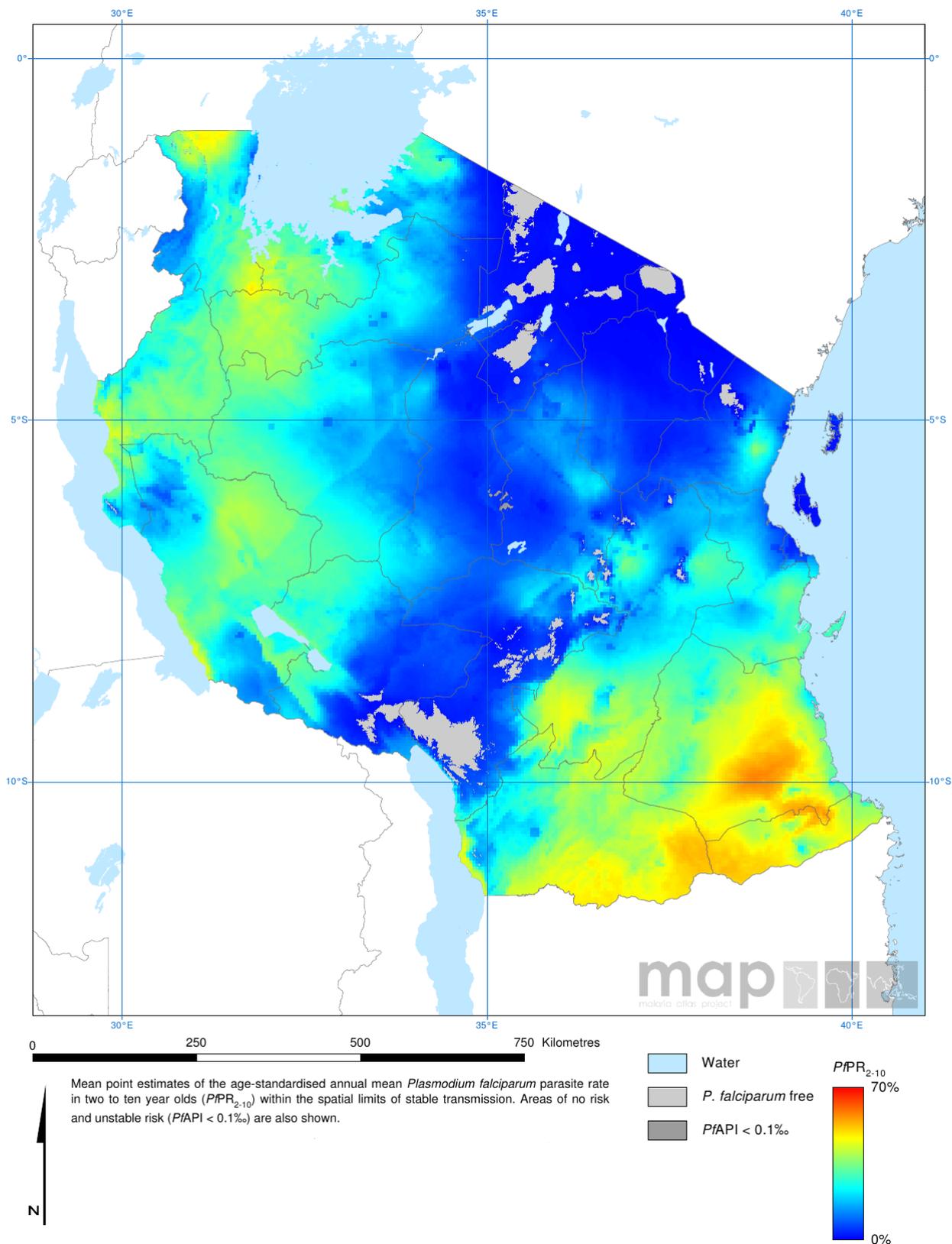
Figure 1.3.:   Spatial Distribution of Malaria Endemicity In Tanzania (United Republic of), 2010

Adapted from Gething et al. (2011)

2. Introduction

The last chapter presents the overview of this research with the aim of providing a quick understanding of every aspect of this work. In this chapter, we render the background to this research, the problem statement, the objectives, the justification for this research, the geographical description of our study areas, and, the epidemiology of malaria.

2.1. Background

In the year 2000 Africa's heads of government convened in Abuja, Nigeria's capital to chart a new course for the control and possible eradication of malaria from the continent. At the end of the conference they came out with a communiqué otherwise called "The Abuja Declaration", of which the main target was the halving of malaria by the year 2010. This target falls in line with the fifth Millennium Development Goal (MDG) which also has as part of its objectives the halving of malaria prevalence and incidence by the year 2015. One of the measures proposed in the "Abuja Declaration" is the reform of the way each country's National Malaria Control Programme carries out its operation. The document was signed by all the 44 participating African countries.

Insecticide treated nets distribution, indoor residual spraying, distribution of free anti-malaria drugs, and maintaining a clean environment are part of the measures that the Roll Back Malaria Partnership intend to use to achieve the aims and objectives of the "Abuja Declaration". Though the use of nets to prevent nuisance bites from mosquitoes has been in use for long but they have not been effective in preventing/controlling the disease as they only serve as physical barriers against the vector. The insecticide treated nets are bed nets impregnated with pyrethroid chemicals and serve not only as physical barrier between man and the vector but also has an "excite-repellent" effect which serves as a chemical barrier between man and his host (WHO, 2007).

The World Health Organisation (2007) divides the Insecticide Treated Nets into two: The Conventionally treated tets and the long-lasting insecticide nets. Whereas the conventionally treated nets needs to be treated with insecticides after every three washes or at least once in a year, the long-lasting insecticidal nets are treated with insecticides directly from the factory and would not need to be treated until after 20 washes or three years after its first use.

Since the penetration of the insecticide treated nets into the African market its coverage on the continent has been very slow with less than 5% of children sleeping under an insecticide treated net while nearly 20% sleep under any form of nets (Hill et al. 2006).

We mentioned in the previous chapter that malaria is a disease that has been with man since time immemorial and it is probably one of the first set of diseases that man has to deal with. The disease was eliminated in Europe and the United States in the 1940s where the insecticides dichloro-diphenyl-trichloroethane (DDT) and dieldrin (DLD) were used in the indoor spraying of the households.

Quinine was the first drug that was used in treating the disease - even before the advent of insecticides - it is an extract from the bark of a Peruvian tree called Cinchona and was discovered in the early seventeenth century. Chloroquine was discovered in the early twentieth century, which shares certain chemical properties with the drug Quinine. These drugs were quite effective in the control of the disease, but the parasite developed certain resistance to them after some time. At the moment the Artemisinin-Combination-Based therapy (ACT) of Artemisinin, Artemether and Lumefantrine are being used as the first line treatment for uncomplicated Plasmodium falciparum malaria nevertheless the Quinine based therapy is still being used as first line treatment for *Plasmodium vivax* and *Plasmodium ovale* malaria (WHO, 2007). In Africa the Quinine-based drugs are still very popular among households because of their price and uncomplicated dosage.

Malaria being a health problem could prevent members of the farming household from achieving their maximum productive capacity. Apart from the fact that it has the ability of imposing unbudgeted costs on the households' income it could impose some social costs on the household also. At the 'Abuja declaration' summit, African heads of government emphasized that malaria costs the African government about twelve billion United States dollars annually and could account for about 25% of a poor household's income (WHO, 2003a).

In summary, malaria imposes unexpected health, social and economic costs on the farming household by reducing the farmers' productive capabilities. We believe that for Africa to take its place in the comity of nations its farming households' productive capabilities must be enhanced through the maintenance of good health through a proper malaria control programme. This research aims to quantify the amount of money the household is willing to pay for malaria abatement through the purchase of whatever prophylactic measures that is at the disposal of the households in question. The question of the validity of this measure raises considerable scope for nuanced econometric investigation and we motivate this in the next section.

2.2. Problem Statement

More than two billion people live at risk of contracting malaria in the world with over 300 million clinical cases of the disease per year (Keiser et al. 2005). A great percentage of this occurs in Africa with the continent, accounting for nine out of every ten cases of the disease in the world (Ajani and Ugwu 2008; Breman et al. 2004; Keiser et al. 2005; Samba 2001).

Nigeria; one of the countries we investigate; accounts for 25% of the total malaria deaths in Africa and it is the highest on the continent (Sofola 2008). Over the period 2001 to 2007, the number of people who died from malaria in Nigeria increased from 4,317 to 10,289. Tanzania records about 11 million clinical cases per year (Tanzania Ministry of Health and Social Welfare 2010) and it constitutes about 14% of outpatient visits in Ethiopia (USAID and The Ethiopian Government, 2011). It is not, however, a story of doom and gloom for Africa's fight against malaria, for example, Hay et al. (2005) and Uguru et al. (2009) observed that there was a decrease of almost 50% noticed in some African countries like Rwanda, Eritrea and [Mozambique]. Also, Tanzania has also noticed an improvement in the distribution of insecticide treated bed nets and malaria prevalence reduced from 20% to 18%.

On the macroeconomic scale, there seems to be a correlation between malaria, productivity and economic growth (Sachs and Malaney 2002 and Malaney et al. 2004) and this could explain why governments (with the aid of international partners) of malaria endemic countries expend a huge part of their annual budget on the control and/or eradication of the disease. We remind the reader that our research is a microeconomic study, part of which is enquiring into the correlation between malaria and productivity. The malaria eradication programme in the 19th century in the Mediterranean and some parts of Asia coincided with the period of industrial and economic growth in these areas (McCarthy et al. 2000 and see Acemoglu et al. 2001 for more elucidation on this). In Africa, the regions of the extreme North and South which are low malaria incidence areas have also been discovered to be the most developed parts of the continent (Gallup and Sachs 2001). Sub–Saharan Africa, which has the highest cases of malaria in the world, is part of the least developed regions of the world.

Though, agriculture employs over 70% of the labour force and contributes a great percentage to Gross Domestic Product, it is surprising that malnutrition and hunger are still persistent in Africa. "For example, between 20 and 75 percent of the population in 29 countries in Central, Western, Eastern and Southern Africa were reported to be undernourished" (UNEP 2007). Thus increasing agricultural productivity is *sine qua non* for Africa's survival. In fact, increasing agricultural productivity forms a great part of the policy strategy and manifestos of most African governments, but it seems attaining this goal is elusive.

In order to achieve food security and avoid undernourishment we believe a study of the impact of malaria on the microeconomic level could be of help in finding a solution to this problem. Malaria affects labour mobility, mortality, investment and fertility decisions of the household the result of which could be efficiency and productivity losses to the household (Malaney et al. 2004). There seems to be a consensus among several researchers that malaria affects the "poorest of the poor" households and it poses a considerable economic burden on the society (Ajani and Ugwu 2008; Chuma et al. 2010; Ettling et al. 1994; Jimoh et al. 2007; Malaney et al. 2004; Mills 1992; Onwujekwe et al. 2004).

Measuring the impact of malaria is a somewhat delicate process because unlike other diseases like HIV/AIDS that have similar effects across age ranges malaria's effect across different age-range is quite varied. For example, its impact is mostly described as being a major cause of death in children less than five years and in pregnant mothers while in older children and adults the impact is mostly described in terms of morbidity and incapacitation. Also, Ajetunmobi et al. (2012) explained that it is described differently after clinical diagnosis as uncomplicated, severe, or, complicated. This might be the reasons why it is quite difficult for policy makers to deal with the disease. We look at the impact of the disease from the point of view of morbidity and incapacitation of household labour.

Also, the disease tends to manifest itself at the peak of the planting season where labour demand is quite high and at a time where hired labour might not be a perfect substitute for family labour for farm activities (Nur 1993). In other words, malaria prevalence is seasonal in nature. Breman et al. (2001) put this succinctly by emphasizing that a good estimate of the burden of the disease should take into consideration parasitological, clinical, epidemiological and entomological indicators. The way in which our malaria data were collected implicitly took all of these into consideration and we believe this process further serves as one of the merits of this research.

In furtherance of the Breman et al. (2001) assertions, the epidemiology of the disease varies from one region (even from one district to the other) to the other in a country and also from one country to the other. In other words, spatial differences exist in the distribution of the disease within and between countries. For example, according to the Tanzania Ministry of Health and Social Welfare (2010) Rufiji and Ifakara districts recorded the highest malaria prevalence among all the districts in Tanzania. *Therefore, in measuring the burden due to the disease absolute care should be taken in order not to overgeneralize the results.* The malaria data we use in this research was collected on spatial basis and took the spatial differences in the spread of the disease into consideration.

Our main goal in this research is to be able to objectively estimate the Willingness-to-Pay for malaria prophylactic measures in three African countries namely Nigeria, Ethiopia, and, Tanzania respectively. We specifcally raise the following research questions:

- Does malaria impact significantly on the efficiency of the farmers in each of the study

areas?

• How much are the households in each country of study willing to pay for malaria abatement?

• Can we place a value on the amount the households are willing to pay in the future?

We are therefore confident enough to draw-up the following objectives in the next section.

2.3. Research Objectives

Our main objective is to estimate the Willingness–to–Pay for malaria abatement in Nigeria, Ethiopia and Tanzania through the assessment of the technical efficiency and productivity of the farmers in these countries.

The specific objectives are:

I . To estimate the technical efficiency of farmers due to malaria in our study areas

II. To calculate how much the households are willing to pay for malaria abatement in our study areas

III. To make predictive inference on the future amount the households in our study areas are willing to pay

We have presented our objectives, in the next section, we attempt to justify this research with focus on these objectives.

2.4. Justification

Our focus in this research is to arrive at a reliable measure of the amount the household is willing to pay for preventive measures in Nigeria, Ethiopia, and, Tanzania. Also, we analyse the impact of malaria on farming households' technical efficiency and arrive at an index of technical performance of individual households in each of the countries of concern.

Good health and productive agriculture are important in the economy of a nation in the fight against poverty and in the development of its human capital as health capital is an indispensable input in agricultural production (Ajani and Ugwu 2008). Information on the impact of malaria on farming households' efficiency is important so as to target intervention efficiently and equitably and to justify investment in research and control (Chima et al. 2003). The achievement of the Sustainable Development Goals and the Roll Back Malaria objectives are pertinent for policy makers, the continual assessment of

the level of prevalence on country and regional bases is therefore a first step in assessing the malaria burden in Africa and improving control programmes.

Onwujekwe et al. (2004) posits that some researchers advocate for free distribution of preventive measures like the insecticide treated nets while others believe placing a charge on such measures would be a better option. However, the fact remains that the cost of control of malaria is becoming too expensive for government, especially when malaria programmes in each affected African country has to compete with other government programmes for funds. Involving the people by asking them to pay a certain charge will be a way out of this dilemma. For example, as at 2008 in Nigeria over 12 million insecticide treated nets were distributed and about half of these were distributed by commercial means (WHO 2012). In arriving at their position, most advocates of a charge on preventive measures usually use the *Contingent Valuation* technique in arriving at their assertion, but the problem with the technique is its subjective nature because the data is based on the individual's personal feelings, perception, tastes or opinions.

Being subjective means the respondents may not state their true position or feign ignorance on a particular subject matter (Hajek 2012); on the contrary, we make use of an objective technique which is built on the *Compensated Variation* foundation. Under our technique we marry the theory of the household model with statistical inference which results in estimates of the willingness-to- pay for malaria abatement.

We make use of the Bayesian method of statistical inference in obtaining these estimates and in doing this we vigorously employ the Gibbs and Metropolis-Hastings Markov chain Monte Carlo (MCMC) techniques of statistical sampling. Hence our work forms part of the few body of literature in this area, we believe this is another major motivation for our work. With the use of the MCMC techniques we are also able to come up with estimates of what the household would be willing to pay in the future by carrying out predictive inferences on our data. In addition, we take a journey into formal model selection by arriving at a method that enables us to arrive at the most preferred estimates for our data.

The last few sections motivate our research problem and objectives. It also succeeds in motivating the justification of our research. In the next section, we render a geographical description of Africa and our study areas.

2.5. Geographical Description of Africa

In this section we give a brief but detailed description of Africa and its environmental characteristics (Brown and LeVasseur 2006 and Healy 2012 are the main source of information for this section). We also give a concise description of the country profile of our

individual study areas which are Nigeria, Ethiopia and Tanzania. Please see figure (2.1) below for a detailed map of Africa.

Africa is the second largest continent in the world and it makes up one-fifth of the world's land surface. It is located in all the four hemispheres - western, eastern, northern and southern – which influence the type of climate present in Africa (Brown and LeVasseur 2006). The prime meridian runs vertically through Africa while the equator runs horizontally through it (Mrdonn 2012). Africa extends 4,800 miles north to south and 4,500 miles from the longest width east to west (Healy 2012). It is three times the size of the United States and to further emphasize its size it is bigger than a combination of the United States, Europe, India, China, Argentina and New Zealand (Healy 2012).

Africa does not have major mountain ranges, although it has pockets of volcanic mountain adorning different parts of the continent. It is also home to a lot of great lakes examples are Lakes Turkana in northern Kenya, Lake Albert in western Uganda, Lake Victoria, and Lake Tanganyika in Tanzania and Lake Chad (Healy 2012). It is home to the great rivers Nile, Niger, Zambezi, and, Congo. Africa has two deserts; the Sahara desert with an area of 3,500,000 square miles-the biggest in the world- and the Kalahari Desert (Healy 2012). Another name for Africa is a Plateau continent. A plateau is a wide level area of high elevation; most of Africa is above 3000 feet (Healy 2012), this embellishes Africa with its escarpment features and its lack of natural harbours. This, along with the high disease incidence prevents its early discovery by explorers.

Since the equator passes through Africa, countries situated in this location have a very good supply of precipitation and warm weather all year round because of the low latitude. As you move to higher latitude precipitation reduces and drought starts setting in such that at 30 degrees latitude is the desert (Brown and LeVasseur 2006). According to Brown and LeVasseur (2006) the following biomes exist in Africa: "Equatorial rain forest—wet all year, Dry season deciduous forest—short dry season, Woodland—longer dry season, savannah—dry season long enough that the trees are replaced with grasses as the dominant life-form, Grasslands—long dry season, Desert scrub—dry most or all of the year".

Overall, Africa has medium to poor fertile soils that are prone to erosion and land degradation. With a population of about 1 billion people Africa represents 14 per cent of the world's population. The rate of growth of its population is the fastest in the world this is brought about by an increase in birth and fertility rates and decrease in mortality rate. This growth in population puts a lot of pressure on the soil resulting in further land degradation and nutrient mining. Next, we now give a brief description of our study areas – Nigeria, Ethiopia, Tanzania.

Figure 2.1.:   The Political Map of Africa
(adapted from www.mapsofworld.com accessed Nov. 2013)

2.5.1. Nigeria

Nigeria is the most populous country in Africa with a population of over 140 million people and a population density 156 per square kilometre. It has a surface area of 923,708 square kilometres. It lies on the west coast of Africa; it borders Cameroon in the East, Benin on the West, Niger to the North and on the South by the Atlantic Ocean. The climate differs from the arid area in the North with rainfall of between 600 – 1,000mm and last between three to four months in the South which possesses an average of between 1,300 and 1,800mm (or 2 500mm in the coastal areas) of rainfall and lasts from between 9 to 12 months. Just like the climate there are differences in vegetation cover; the Sahel Savannah found in the far north, Sudan Savannah and/or Guinea Savannah in the central part of Nigeria, the rainforest in the South and Mangrove forest in the coastal areas. About 70% of the population live below the poverty line. Nigeria has three administrative levels: Federal, State and the Local government; the local governments are further divided into councils (see Nigeria Ministry of Health, 2008 document for further elaboration on this part).

2.5.2. Ethiopia

Ethiopia has an area of about 1.25 million square kilometres. The Malaria Prevention and Control Plan (2011) approximate the country's population to be about seventy-three million. Ethiopia is one of the least urbanised countries in the world with about 85% of the populace living in the rural areas. Poverty is rife in Ethiopia about 23.2% of the population get their livelihood from 9% of the land area, resulting in land degradation and over mining of soil nutrients. It is comprised of different geographic terrains ranging from 110 metres below sea level to 4720 metres above sea level. This greatly affects the climate and consequently, Ethiopia exhibits different climatic conditions in different parts of the country with a corresponding difference in the malaria level. Annual rainfall varies proportionally with altitude while mean temperatures vary inversely with altitude.

Certain low land areas, for example, Afar and Somali regions have rainfall of between 40 and 400mm and mean temperature of between 25^0C and 30^0C. This is an arid region and in this area the occurrence of the disease is noticed in parts that are close to water bodies apart from this, malaria cases are extremely low. Areas with 400 – 800mm of rainfall are semi-arid regions and are regions of unstable seasonal malaria. Lowland areas with a general altitude less than 1500mm are hot and humid, and, usually receive annual rainfall of between 800 metres and 1200 metres with a mean temperature between 25^0C and 30^0C. As a result of the favourable climatic conditions, malaria in this region is relatively stable.

Highland areas located between 1750 metres and 2000 metres are areas of high rainfall with rainfall values ranging from 1200 – 1600mm with mean monthly temperatures ranging

from 18^0C to 25^0C. These are malaria epidemic areas. Highland areas located between 2000 and 2500 metres have cool temperate climates with annual rainfall ranging between 1200 to 2200mm and the mean temperature is between 15^0C and 18^0C and these are areas of low malaria incidence. Areas located above 2500 metres altitude have a very temperate climate with an annual rainfall of over 1200mm and mean monthly temperature between 4 and 15^0C.

Ethiopia runs a federal system of government with power decentralised from the federal government, down to the regions, then, Woredas (districts) and then Kebeles (localities). According to the Ethiopian Ministry of Health (2006), 68% of the total population live in malarious zones (see The Ethiopian Ministry of Health, 2006 document for further information).

2.5.3. Tanzania

Tanzania; officially referred to as The United Republic of Tanzania is located between longitudes 28^0E and 40^0E, and, between latitudes 1^0S and 12^0S. It has a total land area of 886, 039 km^2 and the water bodies make up $64,131 km^2$, hence, Tanzania has a total area of $947,480 km^2$. It is composed of Mainland Tanzania and the Islands of Zanzibar. The administrative structure in Tanzania is divided in descending order, thus: regions, councils, wards.

The council is the most important administrative authority for public services. Tanzania has two rainy seasons: short rains also called "vuli" are from November to December while the long rains also called "Masika" are from March to May. Tanzania has got four different topographical zones which mean they have different climatic features and this impact on their malaria prevalence rate. The four zones are the coastal lowlands, the central plateau, the basins around lakes Victoria and Tanganyika, and, the highland areas surrounding mount Kilimanjaro.

Tanzania with a population of approximately thirty-nine million people and an annual growth rate of 2.8% is classified as a low income country with a high level of poverty. Tanzania has the third biggest population that has the potential of contracting the disease in Africa after Nigeria and Democratic Republic of Congo (see Tanzania NMCP, 2008 for further elucidation).

The last section succeeds in presenting a geographical description of Africa. Next, we present the epidemiology of malaria with the sole aim of understanding how the disease spreads in Africa and the lifecycle of the vector, mosquito.

2.6. Epidemiology of Malaria

Malaria epidemiology has to do with the way malaria is spread in a particular population and location. This involves the study of the parasite (*Plasmodium* species) - vector (mosquito) – host (man) cycle. A further understanding of this we believe will help the reader understand the significance of the disease and probably understand the problem of developing an effective vaccine to prevent this disease by researchers. Before we start, we would like to define some terms which would further improve the understanding of the reader:

Incidence: is defined as the total number of cases reported through the public health system.

Prevalence: is the total number of cases in a particular area, incidence included. In other words incidence does not include malaria treated at home and/or at local herbalist places and so on while prevalence includes all of these

Endemic Malaria: This is a situation where stable malaria is a regular occurrence in an area year-in, year-out over a long period of time. Endemic malaria only happens in stable malaria zones.

Epidemic Malaria: This is a sudden marked increase in malaria incidence in an area. This happens in unstable malaria zones.

Stable Malaria: Here high malaria transmission rate is steady over time, though, seasonal fluctuation may occur.

Unstable Malaria: Here the amount of transmission is not steady but a sudden increase in transmission results in an epidemic. Those who live in stable malaria zones are less likely to have severe episodes or die of the disease when compared with those from the unstable malaria zones.

Malaria is a vector-borne disease caused by the parasite of the genus *Plasmodium*. It is usually transmitted through the bite of an infected female mosquito of the genus *Anopheles*. The *Anopheles gambiae* group and the *Anopheles funestus* group are the most common vector of malaria in Africa. In Nigeria, for example, they are responsible for 18 to 145 bites per person per year and 12 to 54 bites per person per year respectively. Other species like the *Anopheles moucheti, Anopheles nili, Anopheles pharaoensis, Anopheles longipalpis, Anopheles coustani and Anopheles hancocki* play a minor role in the transmission of the disease in Africa. In Tanzania, 80% of the population live in regions of stable malaria while the remaining 20% live in regions of unstable malaria (Tanzania NMCP2008).

Four species of *Plasmodium* are known to cause malaria. These species are *Plasmodium falciparum, Plasmodium malariae, Plasmodium ovale, Plasmodium vivax. Plasmodium*

falciparium is, however, the most common and virulent of these and it is present in virtually all parts of Africa. It is associated with significant morbidity and mortality. Other species – *Plasmodium malariae* and *Plasmodium ovale* - make up about 2% of malaria cases (Tanzania NMCP2008, Nigeria NMCP2008, Ethiopia NMCP2006). However, there are major differences in the type of the minor parasites that cause malaria in Africa, for example*, Plasmodium vivax* is not found among indigenous Nigerians (Nigeria Ministry of Health 2008) while *Plasmodium ovale* has never been reported in Ethiopia (Ethiopia Ministry of Health 2006).

Most *Plasmodium* species have the ability to remain in the blood stream for a long time (sometimes up to 52 years after the last exposure of the individual to *Plasmodium malariae*) after a poorly treated malaria case even when the individual is not bitten by an infected mosquito. Consequently, *Plasmodium falciparum* could be a major cause of recurring malaria in an individual even when the individual has left a malarious zone for a long time. Malaria is characterised clinically by fever; other symptoms may include headache, chills, or rigours, general weakness, vomiting, loss of appetite and profuse sweating.

The question is; how did the *Plasmodium falciparum* become the leading cause of malaria in Africa? Webb (2011) suggests an answer by saying that long time ago *Plasmodium vivax* was the leading cause of malaria in Africa, however, there was a wide haemoglobin mutation known as *Duffy Red Blood Cell Antigen Negativity* or *Duffy Negativity*. Whoever has this 'Duffy negativity' has a kind of genetic protection against the *Plasmodium vivax* malaria. However, with the continual expansion of the *Bantu* tribes (5000-4000 *B.C* and $1500 - 500$ *B.C*), the change in agricultural practices, and, the types of crops grown; they were exposed to *Anopheles gambiae* which carries the *falciparum* - type of malaria. Unfortunately*, Plasmodium falciparum* did not 'favour' the 'Duffy negative' haemoglobin and rather it causes other genetic mutation such as the sickle cell and the various thalassemia (inherited recessive blood disorders) related diseases. This time around this genetic mutation has its own negative health outcomes; with the carrier of the sickle cell trait from both parents having high protection against malaria and an individual with full sickle traits becoming highly prone to the disease. The net result is a high malaria population in Africa.

The description below of the life cycle and the pathogenesis of the malaria parasite follows Wiser (2010) illustration (Please see figure 2.2 and 2.3 for a diagrammatic representation of this). The life cycle of malaria parasites passes through a sequence of three different types of reproductive stages: a) a single run of sexual reproduction, called the sporogonic cycle, taking place in the Anopheles host; b) a single run of asexual reproduction, called the *pre-erythrocytic* cycle, in a liver cell of the human host; and c) an indefinite number of runs of asexual reproduction, called the *erythrocytic cycle*, in the red blood cells of the human host. Throughout the erythrocytic cycle some parasites differentiate into male and female gametocytes, which, if taken in with the blood meal of an Anopheles mosquito,

will initiate the sporogonic cycle.
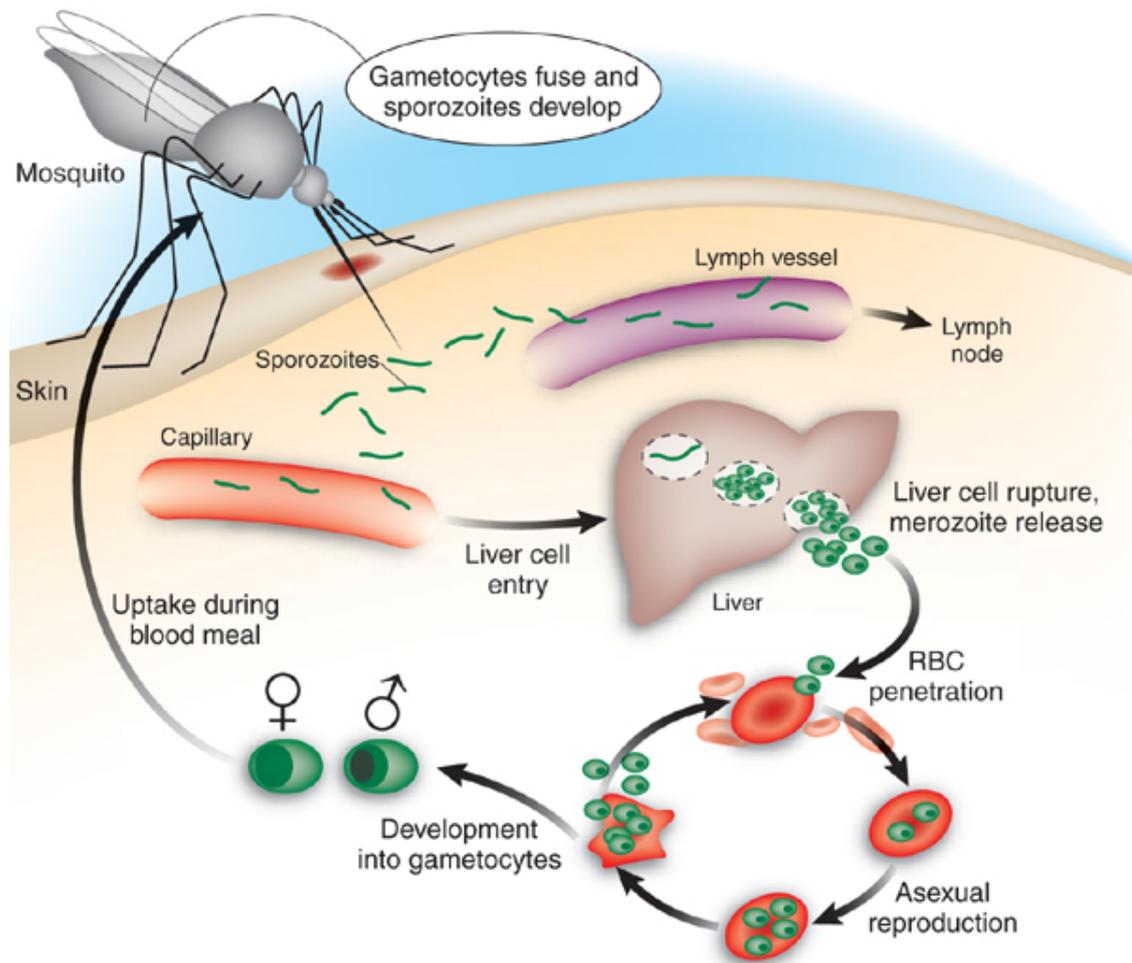
Figure 2.2.:   Lifecycle of the Malaria Parasite
(adapted from Medic19.blogspot.co.uk)
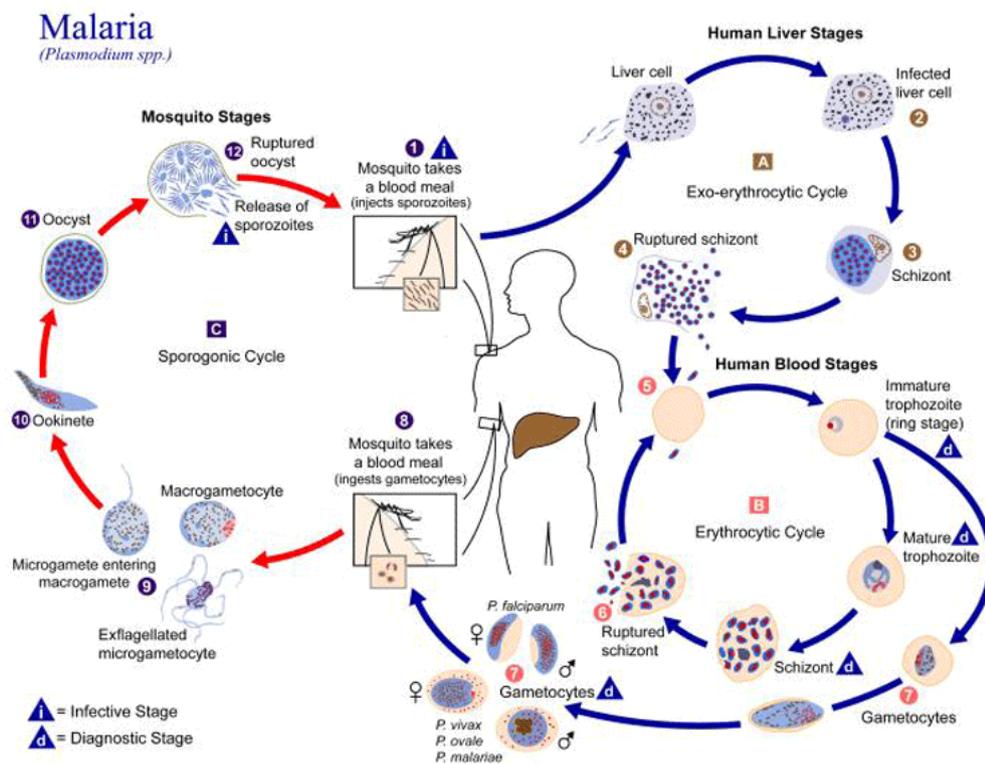
Figure 2.3.:   Pathogenesis of Malaria

(adapted from Medic19.blogspot.co.uk)
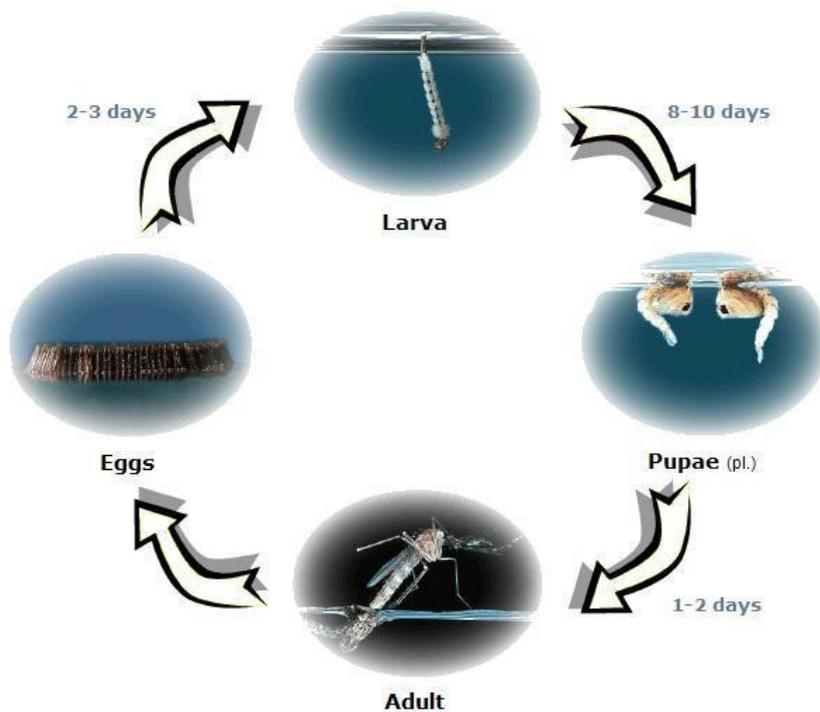
Figure 2.4.:   The Life Cycle of a Mosquito

adapted from *www.vtaide.com* accessed Nov. 2013

The *sporogonic cycle* takes between 9 and 30 days or longer depending on the parasite species but mostly on the temperature. The gametocytes present in the blood meal mature in the stomach of the mosquito. After fertilization, they produce a mobile egg that penetrates and encysts in the stomach wall, where it divides into about 1,000 mobile *sporozoites*. These sporozoites then burst into the mosquito's body cavity and invade the salivary glands, where they are ready to infect a human host in each successive bite.

The habitat of the immature *Anopheles* mosquito is water (see figure 2.4). Eggs are laid on the edge of the water and hatch in 2-3 days to produce larvae (wrigglers), which develop through four larval and one pupal aquatic stage to produce adult flying mosquitoes. Only the female mosquito bites; as it requires blood for the maturation of the eggs; the male feeds on vegetable juices.

The geographic location of Africa provides it with a suitable tropical climate for malaria transmission. The disease is endemic throughout Sub-Saharan Africa, though there could be some seasonal variations in transmission from one country to the other. However, areas in Africa, which have altitude above 1600 meters are not considered as endemic areas, these are areas around the Atlas mountains, Tunisia, and, Libya. Other countries in North Africa like Algeria, Egypt, and, Morocco only have vestigial malaria transmission and infrequent imported cases.

The next section presents the summary of this chapter.

2.7. Summary

In this chapter, we presented the motivation for our study, provided some information on the study area, and, why malaria is prevalent in Africa. We also defined some terms in the literature that are used to describe the recalcitrant malaria/mosquito issue in Africa and highlight the lifecycle of the vector, mosquito, and how it aids the spread of malaria. In the next chapter, we highlight and discuss salient literature relevant to our study.

3. Models And Empirical Evidence In The Household Model

In Chapters one and two, we set in motion the purpose for this research and explain the essence of quantifying the amount the household is willing to pay for abating malaria both in the present and in the future. We go a step further by reviewing extant literature in these areas. In this chapter we will review the literature on the household production model and in chapter five, we review the literature on the composed error model. Thus, presently, we highlight the different thematic developments, and diverse areas the household model have been applied; we also emphasize areas where our research stands vis-à-vis other research in the area.

3.1. Thematic/Methodological Innovations

The household model shows the relation between the farming household as a producer, consumer, supplier of labour and a user of inputs. It tries to show how the behaviour of the household as a producer affect its decisions as a consumer and a supplier of inputs with the sole aim of helping policy interventions (Singh et al. 1986).

All of the models that we survey in this section either bears relation to or are extensions of the seminal work of Singh et al. (1986, pp. 17 - 19). They consummate the several works done on the agricultural households into an eclectic model, which they referred to as "The Basic [Household] model". We present a prototype of the model below:

$$Max\,(U) = U(X_a, X_m, X_l) \tag{3.1}$$

subject to:

$$p_m X_m = p_a(Q - X_a) - \omega(L - F) \tag{3.2}$$

$$X_l + F = T \tag{3.3}$$

$$Q = Q(L, A) \tag{3.4}$$

where: U is the utility function to be maximised which is positive on the real line of numbers, with the commodities X_a, X_m, X_l being agricultural staples, market purchased good, and leisure respectively; they are also assumed to be positive, real numbers.

The quantity Q denotes household production of staples given by equation (3.4), while, ω, L, T, and, A are the market wage for labour, the total labour input, family labour, land, and total available time. All these quantities are positive real line numbers. The quantities $p_m \geq 0$ and $p_a > 0$ are the prices of market purchased goods and agricultural staples respectively.

Assuming the household maximises its utility function subject to a cash-income constraint, equation (3.2), a time constraint, equation (3.3); and the prevailing production technology, equation (3.4). One can rewrite equation (3.3) as $F = T - X_l$ and substitute this and equation (3.4) into equation (3.2), so that:

$$p_m X_m = p_a \left\{ (Q(L, A) - X_a) \right\} - \omega \left[L - (T - X_l) \right] \tag{3.5}$$

By expanding equation (3.5), we have:

$$p_m X_m = p_a \left\{ (Q(L, A)) \right\} - p_a X_a - \omega L + \omega T - \omega X_l \tag{3.6}$$

On further simplification of equation (3.6) and collecting the $X's$ and their coefficients on one side of the equation, we arrive at a single constraint equation of the form:

$$p_m X_m + p_a X_a + \omega X_l = \omega T + p_a \{ Q(L, A) \} - \omega L \tag{3.7}$$

where: $\quad p_a \{ Q(L, A) \} - \omega L$ is a measure of farm profit defined as:

$$\pi = p_a \{ Q(L, A) \} - \omega L \tag{3.8}$$

Note that equation (3.8) only holds if the market for commodities is complete, in other words, the market is perfectly compevtive.

Equations (3.1) and (3.8) are the foundation of the household model (Singh, 1986, p. 18) and it is the subject of several arguments in the literature.

The agricultural household model forms the foundation of our theoretical framework, which we motivate further in chapter four. It is worthy of note that this model does not apply to commercial farms who neither consume a large percentage of their output nor supply a great number of their inputs (Taylor and Adelman 2003).

One issue that generates a lot of argument in the literature is how to deal with production, consumption, and, labour decisions of the household - if they should be treated separately or not - in other words, household decisions may be recursive (that is, production decisions are taken independently of consumption and labour decisions) or non-recursive (that is, production, consumption, and, labour decisions are taken jointly). We are quick to note that, in this research we do not focus on these nuances, but our priority is on how to arrive at a reliable estimate of the willingness-to-pay for malaria abatement using the household model.

Early works in this area commences with A.V. Chayanov in the early 1920's (see Thorner et al. 1966) who attempts to analyse peasant household behaviour in Russia, Krishna (1964), and, Jorgenson and Lau (1969) proposing that production decisions of the household are taken independently of its consumption and labour-supply decisions, in other words, they propose a separable model. In reality, consumption and labour supply decisions are not actually taken independently of the production decisions of the household, that is, these three decisions are taken simultaneously.

The advent of the seminal work of Becker (1965) (Mincer et al. 1963, Dean 1965 and an unpublished work, Owen 1964 cited in Becker 1965 impels his work) spurs a lot of development in this area. He introduces time into the household model, where the total stock of time (ωT) available to a household is divided between leisure, on-farm and/or off-farm production.

The literature witnesses multifarious advancements in the use and modification of the household model since the publication of this work. However, one work that does not agree with Becker's is Pollak and Wachter (1975). They criticise the inclusion of Becker's allocation of time into the household production function. They see time as another commodity with its own utility and demand. They further state that the household time is both an input and output (in the household) devoted to an activity. They observe that the inclusion of time suggests joint production in the activities of the household, hence, the strong assumptions of constant returns to scale and thus the joint production assumption made by Becker breaks down. Also, they reject the entire concept of shadow-price for commodities. They suggest an alternative for analysing the demand for commodities using the prices of goods rather than the prices of commodities. This paper is also one of the first set of papers that criticises joint production of activities.

The publication of Pollak and Wachter (1975) instigates comments and criticisms from Barnett (1977) (He refers to Becker's model as the "new home economics"). He defends Becker's model and proofs that the joint production assumption of Becker (1965) does not break down when time is included in the household production. Specifically, he defends the "new home economics" on the following grounds: that the model does not break the link between the then existing household production function and the neoclassical theory, and, it does not invalidate tastes and technology in the context of shadow prices. He

proofs his assertion by translating these issues into identification problems using a simultaneous structural equation system (He states Pollak and Wachter estimate a reduced form equation). He confirms that Becker's model has neoclassical properties and it is devoid of any under-identification issues. He, however, agrees with Pollak and Wachter that the model is not devoid of (commodity-consumption quantities) measurement problems, but states that these are not the crux of Pollak and Wachter criticisms and that the measurement problems pre-dates Pollak and Wachter(1975)'s publication.

In their reply to Barnett (1977), Pollak and Wachter (1977) agrees with Barnett that joint production does not preclude identification in the model, however, they state that it is not the core of their criticisms of Becker's model. They insist that with well-behaved preferences and convex sets, joint production confounds tastes and technology under shadow prices. They advice Barnett to do away with the idea of seeing the household production theory like the traditional demand function in which commodities are defined on shadow prices. They accused Barnett of misrepresenting their estimated equation but state that they were not specific about the correct equation system. They recommend an equation which serves as a balance between single equation estimation of commodities using shadow prices and the simultaneous structural equation system. Presently, we discuss advancements and scholarly issues debated in the literature.

Singh et al. (1986 p.48) explore the extent to which separability and non-separability can be justified in the literature and assert that non-separability can only be assumed in an agricultural household model when the market for the household staple is incomplete or when the household consumes all it produces, in other words, the household is at a corner; they derive the comparative statics for a one period non-separable model.

Lopez (1986) states another situation where non-separability in the household model should be assumed in the literature. He opines that off-farm and on-farm household activities are not perfect substitutes; and shows the different ways by which the recursive assumption breaks down. He emphasizes interdependence in household utility and profit maximisation decisions of the household; and tests if the assumption of recursiveness against non-recursiveness is significant. He observes that the cross effects between the production equations and the labour-supply equation is high and posits that interdependent models are better for farm household analysis. He shows that models should always assume that off-farm and on-farm behaviour of the household are different. He concludes that the labour-supply elasticities obtained from the recursive and non-recursive models are statistically significant.

Prior to his 1986 paper, Lopez (1982) publishes a seminal paper where he presents the estimation of a joint household model applying the duality theory. This paper is one of the first sets of papers that attempts to estimate a non-separable model using the duality theory. His approach helps to solve the problem of first order conditions in linking the cost minimization demand equation to the profit maximisation production function. One main

advantage of this method is the specification of more complex functional forms with fewer restrictions on the model to be estimated (before this work the Cobb-Douglas functional form was always assumed because of its ease of use, albeit, it is very restrictive).

Also, there have been debates in the literature on whether peasants are utility maximizers or not (De Janvry et al. 1991 cites Vergopoulos 1978) and if it is necessary to use the optimisation principles of economics to analyse their behaviour (Polanyi 1944; and more succinctly put by Dalton 1961). Authors who believe they are not utility maximizers also argue that markets do not exist for them (Dalton 1961). But Cook (1966) disagrees with these views and refers to their views as "obsolete *anti-market* mentality".

De Janvry et al. (1991) assertions are in concordance with Cook (1966)'s proposition; this motivates them to develop a non-separable household model for peasants under these market conditions. They carry out simulation exercises under different market conditions, including failed markets, specifying a Leontief Profit function and a translog indirect utility function. They recommend policy interventions that can be used to increase peasants response to price incentives and propose programmes that will improve the productivity of peasants and put more money in their hands. Such programmes include infrastructure development, increased competition among local traders, and increased market information.

Jacoby (1993) also introduces a general method of estimating a non-separable household model in a developing economy using the shadow price of wages (or the opportunity cost of time). He opines that his method can also be used for self-employed household members and/or voluntary household workers. The method involves estimating an agricultural production function for all labour used by the household (he uses adult male and female, but his methods can be applied to all other labour types like children and hired labour). The marginal product from this estimation is then used as a substitute for wages in the structural labour supply equations using Ashenfelter and Heckman (1974) methods. He applies this model to the 1985-1986 Peruvian Living Standard Survey data and concludes that rural farmers tend to be Pareto efficient in their allocation of resources.

Taylor and Adelman (2003) ventures into the presentation of another general model of the type of Singh et al. (1986) which they referred to as "The Household-Farm Model"; their choice of this nomenclature, they say, was founded on the use of different empirical models to analyse the household model in rural economies. They develop a "General Algebraic Modelling System" for analysing the impact of government agricultural policies on the rural economy. In illustrating a household-farm model, they depict a two-good household-farm model with the household deriving utility from the consumption of food and leisure. This differs from Singh et al. (1986) model that depicts a three good household-farm model with utilities derived from food consumption, purchased goods, and leisure. They present their model under three market scenarios - zero , perfect neoclassical, and mixed markets and apply this simple household farm model to 1993 village cross-sectional data

of one hundred and ninety-six households in Mexico. Their simulations explain the small impact of Mexico's agricultural policy on variables like output, income, and migration. Part of the limitations of their research were assuming choices and income are shared by the households, and, the non-inclusion of some important exogenous variables.

Apart from separability, the household model was developed under a static equilibrium assumption, Innes (1993) took these assumptions further by developing a household model for a two period planting-harvest season for developing countries and he simulates the effect of government policies on this model. Saha (1994) modifies Innes (1993) household model by including risk and uncertainty in his model. He observes that a reduction in price and yield would lead to increase household consumption and production. He finds that the risk-free model largely underestimates the households' consumption, production and labour supply responses to changes in income and price. As a result, he recommends that policy makers should put in place risk-allaying policy to increase households' consumption and marketed surplus.

Renkow (1990) develops a method of inputting discounted future utility derived from storage of marketable surplus into the household model.and applies it to the International Crops Research Institute for the Semiarid Tropics (ICRISAT) village studies panel survey of between 1976 and 1983. He discovers that the future marketed surplus for households are positive. His conclusion confirms the importance of the research, as he recommends that values of present stock needs to be included in the marketed surplus and demand equations even when these stocks do not respond to changes in commodity prices.

From this section, it is obvious that the *modus operandi* of the household model is quite diverse, hence, we discuss the different utilisation of the model, and, the different data types employed under diverse headings in section 1.2 below.

3.2. Diverse Application In The Literature

3.2.1. Agriculture and Health

The first set of empirical studies (such as Yotopoulos and Lin 1976, and, Kuroda and Yotopoulos 1978) on the household model assumes separability in household decisions. The initial popularity of this assumption in the literature could be due to the fact that the assumption of non-separability comes with a lot of mathematical drudgery which the authors might want to avoid. Also, the assumption of a shadow price for labour and the subsequent calculation of a substitution and profit effect further complicate the derivation of a non-separable household model.

Yotopoulos and Lin (1976) are among the first sets of authors that assume recursiveness in household decision making in production. They attempt to use the normalized profit

function to estimate farm production output in Taiwan. Their research is an extension of their earlier work done in India (see Lau and Yotopoulos 1971). They assume that in the short-run, the households' (firm) price of output and other variable inputs are pre-determined, and are not subject to change by the action of any one farm. In other words, the farm (household) is a price-taker in the produce and labour market. They theorised that production, consumption and labour decisions are only linked through the profit-maximizing behaviour of the household (farm) known as profit effect.

Other researchers - Lau et al. (1978); Kuroda and Yotopoulos (1978) - followed immediately afterwards. Barnum and Squire (1979b) estimate a non-separable household model for rice producing households in Malaysia. This paper seems to lay the foundation for the seminal work of Singh et al. (1986) with Yotopoulos and Lin (1976) motivating their work. However, they claim their work departs from the Yotopoulos et al. (1976) in the area of econometric specification; they utilise the quantity of 'inputs used' in their production function as against the 'price of inputs' used by Yotopoulos and Lin (1976), which varies considerably in cross-sectional data.

Also, Instead of the Logarithmic Linear Expenditure System, Barnum and Squire (1979a) utilize a modified version of this which does not constrain the elasticity of consumption to be unity. They link the production and consumption aspects of the household model through the expenditure system; and observe that the economic cost of rural-urban migration is beneficial to the households in North-West Malaysia because the economic cost of migration is small when compared to the marginal productivity of the migrant labour at his place of origin. They observe that government price intervention on output and marketed surplus is zero. Also, *padi* households obtain considerable benefits from price increases and change in technology and this impacts on the expenditure of the households on non-agricultural goods, which they expect to impact positively on the wage-rate.

Simtowe and Zeller (2006) examine the household model under credit uncertainty for maize farmers in Malawi. They focus on households with varying credit constraints. The authors carry out their model through two different stochastic processes, the first stage involves measuring the probability of being constrained while the second stage involves measuring the impact of access to credit on the two selected groups in the first stage. The authors estimate a probit model in the first stage and a switching regression in the second stage, they estimate these two models using the Cragg (the double-hurdle) model. They observe that different factors cause the two groups to adopt an innovation.

Johansson and Palme (1996) attempts to give an economic interpretation to the effects of unemployment, risk exposure, and, control at the workplace in Sweden under the assumption of unobserved heterogeneity and serial correlation. They estimate a separable model estimating a linear demand function. They utilize data from the 1981 Swedish Level of Living Survey. They generate a sub-sample of male and female in the workplace and find that the male subsample shows negative effect on workplace absence, while, the

female sub-sample fails to meet the Slutsky conditions. Because the female subsample fails to meet the Slutsky conditions, which they state as $(a + a_f f) < 0$ and $(\beta + \beta_f f) \geq 0$ with f being a dummy for female and other variables are parameters to be estimated. They assert that their model could have been misspecified in some respect. Similarly, Johansson and Brafinnafis (1998) use the separable household demand model to assess the economics of workplace absence using the 1981 and 1991 Swedish Level of Living (panel) Survey. This work is an extension of Johansson and Palme(1996)'s work. They apply the generalised methods of moments estimator to their data and report results from the first two moments. They found the elasticities of absence to be -1.8 for males and about -2.7 for females, which they say is in accord with similar studies in the literature.

Elhorst (1994) develops a non-separable agricultural household model to investigate if dairy farming households in the Netherlands take production and consumption decisions simultaneously with particular reference to the milk quota. He draws motivation from the works of Lopez (1984), Furtan et al. (1985), and, Thijssen (1988). One of the high points of this research is the ability to show that the agricultural household model is an extension of the profit-function type household production model. In order to prove this, he compares two different elasticities - one for when production decisions are taken alone and the other when production and consumption decisions are taken jointly. He observes that the difference between these two elasticities is very small, such that the model estimated by a researcher does not really affect the result arrived at significantly. He suggests further research to confirm his findings.

Rosenzweig and Schultz (1983) estimate a household health production function using birth weight as an important input and other variables stipulated in the medical literature. They emphasize that the literature is replete with works on the demand for health inputs (Goldman and Grossman 1978, and, Leibowitz and Friedman 1979), with few exceptions focussing on the technical/biological effects of these inputs on health (Edwards and Grossman 1979 is an exception). These works, they opine, fail to examine the endogeneity of these health inputs, in other words, they assume that the health endowment is constant across the population, and it is not an innate property of an individual. In their case, they examine heterogeneity in health input of an individual using birth weight as an innate property assuming non-separability in health input. They construct both production and a reduced-form demand functions for production of child health and demand for health goods by the household.

They use the 1967, 1968, and 1969 United States National Natality Followback Surveys as the data that meets the requirement of their research. They utilize a two-stage least squares method (2SLS) to estimate their data. In the first stage, they construct a log-linear input demand equation for all the behavioural variables employed in the analysis; substitute the estimates obtained into a translog production function using price, income, and education as instrumental variables to identify health technology. They observe that

the inclusion of heterogeneity in the model conceals the significant positive impact of early pre-natal medi-care on child's health and underestimates the significant negative effect of maternal smoking - a proxy for the health of children - on the rate of foetal growth. They place two important caveats in their work; the first being estimates that may be sensitive to other variables may have been omitted in the analysis, also, the instruments used may correlate with the health endowment variable. The development of the household model over the years has its use in assessing transaction costs of peasant households in agriculture.

Ulimwengu (2009) uses the non-separable household model to investigate the effect of the farmers' health status on agricultural productivity and poverty in Ethiopia. He observes that on the average, agricultural production per unit of input is higher for healthy farmers than for households affected by illness. His research finds that healthy farmers earn 137 birr more per year than those affected by illness. His regression analysis shows that health impediments have negative impacts on agricultural efficiency. He states that investment in the health of rural communities will not only increase efficiency and income, but on the rate of return on other investments like education and extension.

3.2.2. Malaria

The incorporation of malaria into the household model is not common in the literature. As we have pointed out in chapter two, the epidemiology of malaria and its data related problems are one of the reasons why its incorporation in the household model has not been common. Notable among such literature is Badiane and Ulimwengu (2013). They use the non-separable model of Singh et al. (1986) to investigate into the effects of malaria incidence on agricultural efficiency in Uganda. They present a health production function similar to the one used in this study (please see Badiane and Ulimwengu 2013p.16). They use data from the 2005 - 2006 Ugandan National Household Survey in their analysis. They generate an index for malaria index and use this value in their analysis.

They state that the decision to focus on the household in their study was because cost of illness decisions are taken at the household level. Their argument is supported by Alaba and Alaba (2009) and Russell (2004). They observe a monotonically increasing relationship between malaria incidence and agricultural inefficiency. However, their finding suggests that hospital consultation charges have the most impact on individual households' morbidity and efficiency. They also observe that health expenses for males are lower than for females. They argue that households in their study area do not just expend on malaria prevention because of its burden, but rather the households consider malaria prevention as a way of improving agricultural efficiency and income.

3.2.3. Transaction Cost

Omamo (1998) uses a non-separable household model to investigate a situation where trading decisions and transaction costs are endogenous. He analyses both variables in the absence and presence of risks. Interestingly, he observes that in the absence of risk smallholder households consider the issue of transaction cost as paramount in their acceptance of a new innovation or technology. His findings are important and might explain the reason peasant farmers decline certain innovations as most innovation often do not think of the transaction costs to the farmers.

Also, Key et al. (2000) uses the non-separable household model to investigate the effect of transaction costs on the household's supply response. However, unlike Omamo (1998), they remove the role of risks and credits on the household model they construct. They dichotomise transaction costs into fixed and proportional transaction costs, and, test for the significance of these two types of transaction costs. They apply their model to the Mexican "ejido" data collected by the Mexican Ministry of Agrarian Reform in 1994. They observe that both types of transaction costs are important in explaining the behaviour of smallholder households with the proportional transaction cost being more important in making selling decisions than buying decisions. As a result pricing policy will have varying effects in different sub-sectors of the farming population; also aggregate supply will be affected by changes in transaction costs through its influence on market participation. Lofgren and Robinson (1999) introduce transaction costs into a non-separable household model in a mixed complementarity Computable General Equilibrium (CGE) setting to analyse the impact of exogenous changes on African farmers.

3.2.4. Energy

The household model has also been used in assessing demand for energy by the household. Heltberg et al. (2000) examine the demand and supply for firewood in rural India. They estimate a non-separable household model using the maximum entropy approach and conclude that policy makers should ensure that the vital elements causing degradation are tackled otherwise efforts at increasing the number of forests may be unsuccessful. Amacher et al. (1996) follows this assertion in developing the theoretical framework for analysing households' demand for firewood in Nepal. They state that households usually hire labour for agricultural activities, however, for unobservable reasons, they decide to use family labour to gather firewood. The authors further assert that the ability to collect firewood depends on households' valuation of the total time available and its labour-leisure preference. As a result, the household allocates labour for firewood collection based a virtual wage rate that is only known to the household itself. The authors use a combination of the two-stage least squares, the tobit and probit regression in estimating their data.

They obtain data for their analysis from two different surveys, one from the *tarai* region and the other from the mid-hill region of Nepal. Their results show greater firewood scarcity in the mid-hill region, firewood prices are also higher in the mid-hill region. This result, they say, is contrary to the physical evidence on the ground, and they query reliance on physical evidence rather than on economic evidence of scarcity for policy interventions. They observe that households that do not rely on the firewood market for their consumption respond faster to economic measures of scarcity rather than physical measures of scarcity. Also, they opine that there are behavioural differences between households that purchase and households that collect firewood; this distinction, they emphasize, is important during policy making.

Chen et al. (2006) examine the demand for firewood and coal consumption in rural China. This work follows similar works by Amacher et al. (1996) and Heltberg et al. (2000). They use the non-separable household with emphasis on the time allocated to different activities by members of the household. They observe that distance to the forest has a positive correlation with coal consumption but negatively correlated to firewood collection. They also observe that the possession of improved stove (cooker) does not prevent rural households from sourcing energy from the forest. As a result, they opine that the improved stove (cooker) programme in rural China might not be re-visited by the government. They suggest that policies that gear towards education may have a cross-effect of reducing the demand for firewood in rural China.

Another enthralling application of the household model is in the assessment of consumer demand and welfare effects due to changes in government policy on energy or the environment by Brãnnlund and Nordstrãm (2004). They capture the demand for non-durable goods that use energy by households in Sweden using micro and macro-data from the Swedish Household Expenditure data and the Swedish National Accounts. The microdata contains cross-sectional data on expenditure on fuel, public transport, other means of transport that uses fuel, home heating and other non-durable goods, while the macro data is a time series aggregate of all of these variables.

They assume separability in household labour-leisure choice and choice of consumable items; they also test for these assumptions in their study. The authors specify a quadratic Almost Ideal Demand System functional form applying a two stage budgeting model to these data. Their motivation for the study is the Kyoto convention of 1997, which states that the emission of greenhouse gases from the developed countries should be 5% less than the 1990 levels between the periods 2008 - 2012. Their research shows that households in less populated areas carry large amount of the CO_2 tax burden in Sweden. Similarly, low-income households carry a larger share of the tax burden than high-income households.

They justify their assumption of separability in labour supply and household consumption by assuming that their study looks at the short-run scenario and use labour supply as a conditioning good; which is restricted due to involuntary unemployment, and, fixed

working hours for the employed. They, however, caution that these might not be the case in the long-run. Another reason is that, since Sweden is a small open economy, the households are price-takers in the market. They face the limitation of data to effectively capture the environmental effect due to tax changes.

Micklewright (1989) develops a household micro-data from an aggregated data in order to understand the United kingdom's energy demand. He estimates an Almost Ideal Demand system of equations for gas and electricity each under a separable household model framework. He observes that price and income responses depend on the nature of the tenancy agreement and the type of household equipment owned other things being equal. He also opines that the research should be carried out under a non-separable household framework. Mekonnen et al. (1999) utilizes the non-separable agricultural household model to assess the production and consumption of energy in Ethiopia in the presence of market imperfections focussing on firewood as the source of energy. The data used consists of four hundred and nineteen rural households in Ethiopia. They apply the two-stage least squares approach to their model; and recommend policies that will reduce firewood collection time and encourage dung use as manures as part of the policies that could be implemented by policy makers.

3.2.5. Demand and Consumption

Following their work in 1976, Lau et al. (1978) assume separability in household consumption behaviour. They input leisure as one of the commodities that the household has to consume and assume utility maximisation for the consumption behaviour of the household. The authors estimate a Linear Logarithmic Expenditure System consisting of Leisure, agricultural commodities, and non-agricultural commodities using household cross-sectional data from 1967 to 1968 for a province in Taiwan. They observe that the labour supply/leisure decisions of the household depend on the size of the household, wage rate, and price of non-agricultural goods. The classical demand theory assumes that consumption decisions of the household are taken independently of the production and labour supply decisions (see Deaton and Muellbauer 1980 for further explanation).

However, Chavas (2006) postulates that the classical consumer theory breaks down when food intake affects productivity as the theory becomes inappropriate to describe consumer behaviour. He introduces time as a factor into nutrition and food demand and notes that the present food insecurity and malnutrition demands in the world tend to focus on current survival with less attention on future food security. The author argues that this discourages future investment, economic growth, and capital accumulation. He avers that, other things being equal, food transfer may have stronger effect than an equivalent cash transfer.

Browning and Meghir (1991) consider the issue of separability in their analysis of the effect of male and female labour supply on the demand for goods. They utilize a panel data of six years from the United Kingdom's family expenditure survey from 1979 to 1984 and analyse a seven commodity demand system using dummies for hours of work done and participation in the labour market as conditioning variables. They test if the assumption of separability in most demand and consumption literature is correct and also consider the issue of endogeneity in their conditioning variables. Their results show that labour hours and participation decisions are endogenous. The non-inclusion of either of the conditioning variable results in biased estimates. They place a caveat on these results because they believe that these results are conditioned by the underlining hypothesis they made in the research. However, they state emphatically, that the rejection of the separability assumption is the most significant finding of their research.

The versatile use of the household model has seen authors using just a portion of it to motivate their research. For example, Holloway and Ehui (2001) use the cash-income constraint aspect to investigate the consumption and willingness to pay for extension services in developing countries. They investigate the Ethiopian Milk market and observe that households are willing to pay between one and seven Ethiopian Birr for services of an extension agent. Considering the magnitude of the fixed cost value, they state that variable costs cannot preclude the possibility of privatization in the Ethiopian milk market.

Dalton (2004) assumes a separable household model format to generate a hedonic pricing model for the production and consumption of goods. He tests the model against other similar but more restricted models and comments that an augmented model that follows the household model is better. He observes that the failure to include production and consumption attributes results in biased estimates; the study also found that the yield was not a significant factor that determines farmers' willingness to pay for upland rice.

Edmeades (2003) uses the non-separable household model to investigate the determinants of farm-level choice and the demand for a particular variety of banana in Uganda. He uses Heckman's two-step procedure for count data models to analyse his data. He finds that factors that influence the choice and growing of a banana cultivar are cultivar specific. He finds that production decisions are important only during the time of planting and taste is important for consumption decisions.

De Janvry et al. (1992) examine the impact of the introduction of the Structural Adjustment Programme in Morocco. They use a computable non-separable household model in their analysis. While they place a caveat on their observations, they state that higher prices of cereals results in the transfer of resources from animal farming to cereals, they also observe that if prices of livestock increase with that of grains, it will increase the use of child labour in herding and overgrazing. They affirm that this policy has a heterogenous impact across the population and different members of the household. Saha

and Stroud (1994) utilize 1975-76 and 1984-85 International Crops Research Institute for Semiarid Tropics (ICRISAT) panel data from Shirapur in India to investigate a non-separable household model of farm storage under price risk. They apply Renkow (1990)'s model of the household to this research. They observe that storage choices in Shirapur are sensitive to price risk attitudes, but they remark that their model needs testing under situations where price volatility is high.

Doiron and Kalb (2004) examine the demand for formal and informal child care on household labour supply in Australia. They utilize data from the 1996 Australian Child Care Survey. They estimate a two stage model, in the first stage, they estimate a bivariate Tobit model for demand for formal and informal child care, the predicted costs are integrated into the budget constraints, and this is used in generating a discrete choice labour supply model. They estimate a separate model for both couples and single parents. Their results show that increase in the price and cost of child care affects single parents more. However, they observe that the effect of child care on married men is negligible. They conclude that their result is in accord with that obtained in the United States and the United Kingdom.

Also, Strauss (1984) utilizes the elasticities of marketable surpluses of agricultural households in estimating a non-separable household model. The surpluses are both output and labour surpluses of the household using data from 138 households in Sierra-Leone during the 1974 - 1975 cropping year. He uses the quadratic expenditure system to estimate the demand side of the model. On the production side, he employs the constant elasticity of transformation to specify the dependent variable and the Cobb-Douglas production function for the input side of the model. He discovers positive own-price elasticities for expenditure commodities and observes the cross-price elasticities to be substantial and negative in sign. Hence, he emphasizes the significance of price effects on marketed surplus of other commodities.

3.2.6. Labour Supply and Time Allocation

Browning et al. (1985) attempt to generate a framework for analysing consumption and family labour supply under uncertainty using the integrated life cycle models and the profit function. They apply panel data from the British Family Expenditure Survey from 1970 to 1977 to their framework. They observe that their theory explains the life-cycle behaviour for hours worked and consumption separately, but it is incapable of a joint explanation of these two parameters over a business and life-cycle pattern.

Wales and Woodland (1976) examine the household utility function and labour supply response using the separable household model. They emphasize that their work differs from other research published about this time because instead of estimating two reduced

form equations in which the total time worked by each spouse is the explained variable while wage rate and other socio-demographic factors are used as the independent variable, they generate a Cobb-Douglas and Translog indirect utility function using the family's after tax income. From these, they obtain the labour supply curve for the husband and wife and the associated elasticities. Another difference between their work and other works done about this time is that their data consists of households that are constrained by the total number of hours of work they have to do. Their results did not conform to all regularity conditions for utility maximisation, it also shows that the household may not be maximizing utility. They conclude that market responses are the result of positive and negative individual responses.

The issue of separability in the labour market has also been examined in the literature. Benjamin (1992) examines separability as it relates to the household composition, labour supply, and, labour demand. He remarks that in a competitive market, the household is a price-taker, this necessitates the need to separate labour supply from labour demand decisions of the household, he corroborates Singh et al. (1986) proposition. He endeavours to confirm if in the absence of the labour market, the composition of the household is a major factor determining household labour use on the farm. In other words, farm employment is not dependent on the composition of the household. He uses data from rural Java, Indonesia to carry out his analysis and concludes that the evidence is not enough to reject the hypothesis that farm labour allocation decisions are independent of the household structure. Some researchers estimate a non-separable household model based on the fact that households may decide not to hire labour for certain activities in the home and decide to use family labour (for example, Amacher et al. 1996).

Singh et al. (1986, pp. 7) emphasize that this situation causes the recursive property of the household model to breakdown. Daunfeldt and Hellström (2007) analyse the allocation of time to production activities using 1983 and 1994 Swedish household data. They emphasize that the data differs from most household data because it involves an interview of both spouses in the household. One of the econometric issues the authors attempt to solve is the zero recording for the time allocated for household activities. They note that the literature is inundated with the use of the tobit model to handle this situation; which is only used when the choice of activity is from a single stochastic process. They emphasize that the Cragg model (also called the Double-Hurdle model) is better when the choices are from two separate stochastic processes. The likelihood ratio test is used to test the preferred model, and they discovered that the Cragg model is better that the Tobit model; hence, they conclude that time allocation to household activities takes two separate stochastic processes.

Reardon (1997) examines income diversification in the rural household with special focus on the nonfarm labour sector. He finds that the nonfarm labour sector is an important income generating sector for the household, but its distribution is poor; he notes that

this is as a result of entry barriers and market segmentation which might lead to uneven distribution of land and other inputs in the rural area.

Blundell and Walker (1981) do not support separability in households' decisions between leisure and time, and, the allocation of total expenditure on commodities. They emphasize the importance of analysing a joint model, wherein the allocation of time between work and leisure, and allocation of total expenditure between commodities are non-separable. The authors apply a generalised linear expenditure system to the United Kingdom's family expenditure survey of 1974, which is a sample of two adult manual worker families. Because they include wage earning females who participate in the labour force in their analysis, they had to resolve selectivity bias in their analysis. They use Amemiya (1974) and Heckman (1979) techniques to resolve this econometric issue. Their results show strong household composition effects on female labour supply which impacts on the female leisure time and expenditure decisions.

Skoufias (1994) introduces an empirical framework for analysing production and consumption decisions of the household jointly. He uses Jacoby (1993) approach which provides an alternative to analysing a non-separable household model. He introduces shadow wages for male and female labour in India and then obtains marginal productivity values for each gender using a Cobb-Douglas production function. He substitutes these values into the structural model for labour supply; he observes that there is a significant relationship between total time available for work and changes to household income opportunities.

Azzi and Ehrenberg (1975) develop a simple model for the allocation of the household's total lifetime between time for attending religious activities and other activities. They apply their model to a combination of 1,500 United States households who attend church regularly. Their paper also discusses several extensions of their model and also test several hypotheses.

Garcia and Marcuello (2001) develop a theoretical model for a household contribution to charity and the implication of tax incentives received from this. They test different hypotheses in their model, which includes the way in which the household takes decisions on how much to donate. They observe that households take decisions in a step-wise manner. In the first step, they decide if they should make a donation or not and in the second step they decide how much should be donated. They also investigate the several factors that influence donation, which include attributes and gender of the household head, the effect of the donation on the overall household income, and also the type of government tax policy in the areas the household reside.

Fafchamps (1993) attempts to describe labour decisions of the household under uncertainties. he looks at the flexibility of West African farmers in the allocation of time between farming activities and leisure. To deal with the non-linearity of structural model he obtains the full information maximum likelihood estimate for the utility and production functions. He claims that this approach leads to desirable estimates. His results show

that the farmers allocation to leisure smoothing is low, labour elasticity is high, as a result, large farms tend to suffer from labour shortages when the rains are abundant. He concludes by attributing the low level of agricultural labour use in African agriculture as being caused by low labour productivity and poor planning from the farmers which leads to labour shortage for weeding in the absence of labour constraints.

Huffman (1976) examines the total time spent by a husband and wife on their farm, they develop a behavioural model for a multi-product farm firm for assessing the labour time and inputs. He fits a Cobb- Douglas production function on 1964 Agriculture census data for Iowa, North Carolina, and Oklahoma counties. The result shows the underuse of hired labour; optimisation of both farm and off-farm work activities and concludes by supporting the proper funding of the agricultural extension programme.

3.2.7. Bio-economics

The household model is also employed in the area of bio-economics. De Janvry and Sadoulet (1995) employ the computable non-separable household model to simulate the general equilibrium in asset redistribution in rural Mexico. Their results show that asset redistribution favours the low income earners, but are not sure of the direction of absolute income gains. Thus, they conclude that redistributing land is *sine qua non* for poverty reduction in rural Mexico.

Ruben and van Ruijven (2001) generate a bio-economic recursive farm household model for households in Southern Mali. The bio-economic model focuses on sustainable use of land, resource use optimisation and the households general well-being. They apply meta-modelling to the production side of the household model when there is no market failure and when market failure exists. In the first case, households rely on the market for labour and animal traction, while in the second case - a case of non-separability - , households practise less intensive farming and lower their feeding which consequently results in negative profit effects but more sustainable use of land and resources.

This section succeeds in presenting the diverse application of the household model in the literature. It has also shown that the utilisation of this model is still popular in the literature. However, the model has also received several disparaging views in the literature. We discuss these views in the next section below.

3.3. Criticisms of the Household Model

The household model has been criticised for two main reasons; it is individual based rather than group based, and it dichotomises total time into leisure and work. Apps and Rees

(1996) list two different schools of thought that develop in the literature as a result of the first criticisms. The first school of thought are those they refer to as the bargaining approach models, which originates with Brown and Manser (1978), and, McElroy and Horney (1981) with Chiappori (1988) defining a general formulation for this group of models.

McElroy and Horney (1981) assume that a household (husband and wife) takes decisions on money-income and time jointly. They posit that the neo-classical utility maximisation theory assumes that decisions are taken individually, however, what is being modelled is actually *group* decisions as against *individual* decisions. They propose that their use of Nash-bargained household decisions help to emphasize the fact that household decisions are taken by bargaining of each household member. They state that the Nash equilibrium does not collapse to neoclassical demand equilibrium (McElroy and Horney 1981 cite Horny and McElroy; 1980, and, Brown and Manser 1978).

The second school of thought consists of Apps-Rees group. This group believes that even when households are not working they can still exchange household goods for market produced goods in order to achieve Walrasian equilibrium in the market for household goods (See Rees 1979, Apps, 1981 and 1982; Apps and Rees 1988; and Apps and Jones 1986 for further explanation). Apps and Jones (1986) applies the Apps-Rees model to a sample of 1,384 families selected from the 1985/86 Australian income distribution survey sample file. The Apps and Rees model basically states that a member of the household can trade household commodities for market commodities in the household. Their work confirms the empirical application of the Apps -Rees model emphasising the importance of incorporating domestic production into the a single person or multi-person household labour supply model.

Udry (1996) states that the proponents of the household models (collective household model inclusive) always assume that the allocation of household resources is Pareto efficient. However, he finds that in Africa, where agricultural production takes place on different plots of land, female headed households tend to allocate resources less efficiently than male headed households. In addition, he does not agree with the proponents of cooperative bargaining and household Pareto efficiency as well. He recommends that a new way of analysing the household allocation of resources is inevitable.

The next section presents the summary of this chapter.

3.4. Summary

In this chapter, we have been able to review the different uses and development in the household model; we believe that other literature not reviewed falls into one or more of

the above categories. The household model (including the different types) has shown versatile usage in the literature and it will continue to define the understanding of the rural household. We use the Singh et al. (1986) type of household model because Mattila-Wiro (1999) states that the model is still the most developed of all the household models proposed in the literature (see Mattila-Wiro 1999; for a critical review of the different types of household model). Also, the decision to estimate a separable or non-separable model depends on the nature and objective of the study, and, the availability and appropriateness of data to capture these objectives. In the next chapter, we present the conceptual framework to this research, which then leads us to chapter five where we review the contemporary literature on the composed error model.

4. Conceptual Framework

4.1. Introduction

In the last chapter, we perused the literature on the household model which serves as the fulcrum upon which our research depends. In this chapter, we attempt to present to the reader the framework upon which our research is based. This is important for two reasons; apart from introducing the reader to the basic theoretical framework which forms the heart of our research, it also gives a proper introduction into our estimation procedure which we will discuss in Chapter eight. The focal point in this chapter is equation (4.13) which is the equation that consummates our remit in this thesis. This is because with equation (4.13) we are able to arrive at a value for the willingness to pay. The equation consists of the product of the price of staples (p_a), a production function of two inputs, labour (L), and land (A); the stochastic frontier model, θ'_j, and a change in the malaria prevalence value, $\triangle\theta$ (we expand the production function in chapter eight to include more than two inputs).

In the next section, we elaborate on how we arrive at equation (4.13) using the basic household model.

4.2. The Basic Household Production Model

The underlying model upon which we premise our conceptual framework is the formal household production model which has far-reaching foundation in the literature. Its foundation in the literature reaches as far back as Becker (1965), Lau et al. (1978), Barnum and Squire (1979b), Singh et al. (1986); Pitt and Rosenzweig (1985); and, Omamo (1998) just to mention a few (chapter three explains further). However, this is not the centre of our research.

In this chapter, we emphasize the basic household model as stipulated in Singh et al. (1986 pages 17 – 19) and then modify it in order to suit the main objective of this research. It is at this juncture we would want to remind the reader of the main aim of our research.

Our research focuses on arriving at an estimate of the farmers' production efficiency and then using this estimate to obtain "reliable" figures for the households' willingness to pay

for malaria abatement. The reader should note that we would like to assume the simplest form of the household which consists of just one individual which is just the farmer. For ease of understanding, we start with the farmer's production function consisting of two inputs, land and labour (this is the same assumption in Singh et al. 1986 page 18), however, in chapter eight we include more inputs that may affect the farmers' productivity.

The household maximizes its utility based on its home produced goods, x_a, market purchased goods, x_m, and, leisure, x_l; subject to certain constraints (restrictions).

The first constraint is that its expenditure must equal its full income. The household expenditure is the product of the price of the market purchased goods, p_m, and the physical quantity purchased, x_m, written mathematically as $p_m x_m$. Its full income is based on the difference between the household sales, household labour income and household endowment income, m. The household sales are based on the product of the prices of home produced goods, p_a, and its marketed surplus. The marketed surplus is the difference between household production of goods, Q, and the household consumption of these goods, x_a. The household labour income is a product of the market wage, ω, derived by the household from the use of its labour and the difference between the total labour input, L, for production and the family labour input, F. When the difference between the total labour input, L, and the family labour input, F, is positive it means the farmer has to hire labour for farm production and when it is negative, it means family labour is more than the total labour input required and thus it can hire out its labour, which serves as off-farm income to the family. The household endowment income is based on the innate skills, talents and characteristics of the household which can be used as a source of income.

The second constraint includes the fact that the household cannot allocate to itself more than the total available time to either pursue productive activities, F or for leisure, x_l. The third constraint is that the total production technology facing the household is based on its output, Q, which is a product of the function of labour and land inputs and health efficiency input.

With all these restrictions in mind and with the employment of the tools of calculus, we believe we can arrive at a reliable amount that households are willing to pay for malaria prevention. In order to arrive at this estimate we summarise the household production model as explained in the above paragraphs in mathematical terms below (where U is the utility function being maximized by the household, and all other variables are as defined in earlier paragraphs above):

$$\underset{x_a, x_m, x_l}{Max} = U\left(x_a, x_m, x_l\right) \tag{4.1a}$$

subject to:

$$p_m x_m \leq p_a(Q - x_a) - \omega(L - F) - m \tag{4.1b}$$

$$x_l + F \leq T \tag{4.1c}$$

$$Q \leq f(L, A)\theta(\alpha_j H_j) \tag{4.1d}$$

The functions, f and θ have positive values and are the production and (in)efficiency functions respectively. A measure of households' health level is $\alpha_j H_j$, where $j = 1, 2, \cdots n$ households and the parameter α measures the how much the variable H affects the farmer's inefficiency. For this study the 'health efficiency' of a household is dependent mainly on the procurement of a particular prophylactic measure. In other words, other things being equal, we expect households' that possess any prophylactic measure to be healthier than households without one and vice-versa. In this way one may compute the impact of any change in health related information on resources via the parameter, $'\alpha'$ in equation (4.1d):

$$\theta^{'}(H_j) = \alpha_j \tag{4.2}$$

To digress, one may argue that 'health efficiency due to malaria infection' variable, H_j, in equation (4.1d) above should have been introduced into the first part of the production technology equation. In other words, it should be part of the frontier part of the equation rather than the inefficiency part of the function. We state that the literature is not exactly clear on which variable should be part of the production aspect of the equation and which variable should be on the inefficiency part of the equation, and, whether a particular variable can affect both productivity and efficiency at the same time (see Coelli et al. 2005 and the next chapter for further elucidation). Also, Pitt and Rosenzweig (1986) in Singh et al. (1986), Croppenstedt and Muller (2000), and, Ajani and Ugwu (2008) agree with our position when they state that the health of the farmer affects his/her ability to allocate inputs efficiently and the ability to take effective decisions as well. We test all of these positions in our model selection exercise in chapter nine. The results of the model selection show that the malaria variable can be on either part of the composed-error

model or it could be on both. However, in order to elucidate our framework, we have intentionally placed the malaria variable in the inefficiency part of the model.

Another school of thought might want to follow the Pitt and Rosenzweig (1985) approach, in which case, the household <u>general</u> health level is assumed to affect not only the households' productivity, but also, provides <u>direct utility</u> to the household. We have refused to adopt this approach because our focus is not on overall (general) health level. The general health of an individual is not only based on the ability to obtain preventive measures but also on the amount of food consumed, the total amount of time worked, the social environment, and, other factors beyond the control of the farmer. Also, the pathology and epidemiology of the disease will make this approach difficult. In the first place, we would require accurate <u>household level</u> data which as we explained in chapter one and two are difficult to collate and obtain. Also, it will involve us directly dichotomising the amount of <u>utility</u> the farmer will derive from preventing malaria as against several possible health conditions like fever and headaches and not just general health.We reiterate that our approach is similar to Badiane and Ulimwengu (2013).

Furthermore, it might be argued that the only input that is affected by the procurement of prophylactic drugs is the labour input. Adegeye and Ditto(1985 p. 55) answer this question succinctly by stating that some economists define labour as the "sole embodiment of production". Moreso, they state that every input (apart from land which is given by nature) in production draws from labour and as such other inputs can be expressed in terms of labour units. For example, they state that a tractor or a bag of fertiliser can be expressed in terms of man days which is a labour unit. They see management as labour which has been further refined by training, education, and, experience. Consequently, labour is seen as the focal point of all production. Thus, the question of whether investment in prophylactic drugs affects the labour input alone or all input in general should not arise as it is two sides of the same coin, since, whatever affects labour affects all other factors of production.

"Back to now"; in the data chapter the reader would notice that the prevalence of the vector mosquito is used as a proxy for the purchase of any preventive methods by the households'. The good thing about the introduction of this quantity is that health measure is allowed to enter the household model depending on how the researcher wants to define the health status of the household.

The question on the mind of the reader is how then does the household make optimal decisions under the above constraints which bears the paramount question of; how does a measure of the willingness to pay for malaria prevention come to be calculated?

In the next few lines, we answer these questions by the use of the tool of calculus. Doing a few substitutions, equation (4.1) above can further be simplified as:

$$\underbrace{Max}_{x_a, x_m, x_l} = U(x_a, x_m, x_l) \tag{4.3a}$$

subject to:

$$p_m x_m \leq p_a[f(L, A)\theta(\alpha_j H_j) - x_a] - \omega(L - F) - m \tag{4.3b}$$

$$x_l + F \leq T \tag{4.3c}$$

The quantity p_a, p_m are the prices of home produced goods and market purchased goods. The household makes an optimal decision at the point $x_a(z), x_m(z), x_l(z)$. The quantity, z, can be seen as a vector of 'controls' available to the household. These controls differ from one household to the other depending on its orientation, location or variables which the household can only decide upon. The maximum-value function is then of the form:

$$\underbrace{Max}_{x_a, x_m, x_l} = V\{x_a(\mathbf{z_j}), x_m(\mathbf{z_j}), x_l(\mathbf{z_j})\} + \lambda_1(\mathbf{z_j})\{p_m x_m(\mathbf{z_j}) - p_a[f(L, A)\theta(\alpha_j H_j) - x_a] + \omega(L-F) + m\} + \lambda_2(\mathbf{z_j})\{x_l(\mathbf{z_j}) + F - T$$

$$\tag{4.4}$$

The quantities $\lambda_1(\mathbf{z_j})$ and $\lambda_2(\mathbf{z_j})$ are the Lagrange multipliers for each individual constraints, such that $\lambda_1(z_j)$ can be interpreted as the marginal utility of income while we interpret $\lambda_2(z_j)$ as the marginal utility of time (See Chiang and Wainwright 2005 for further explanation). Before we go on, we would like to remind the reader that health is seen as a very important resource to the household in order for it to achieve its maximum productive output. We are now set to answer the second question on how much the household will be willing to pay for malaria mitigation. In doing this we apply the Taylor's series to the indirect utility (objective) function in (4.4) above (with a focus on the quantities m, θ and $\lambda_1(z_j)$. The function $\lambda_2(z_j)$ disappears in the equation as it will give zero when it is maximized. With the subsequent application of the *Envelope Theorem* and the incorporation of the *Roy's Identity*, we are able to arrive at a reliable amount that the household is willing to pay for malaria abatement. Thus, we digress a bit, in order to explain what the Taylor's series is all about.

The Taylor's series is used to expand a continuous function up to the nth order or infinity and the coefficients of each variable in the expansion may be expressed in terms of derivatives (Chiang, 2005); it makes expansion of a function beyond the first order easier.

The Taylor's series expansion about an arbitrary point, x_0 is stated in generic form as:

$$f(x) = \sum_{n=0}^{n} \frac{f^n(x_0)}{n!} (x - x_0)^n + R_{n+1}(x) \qquad (4.5)$$

Where $f(x)$ possesses continuous derivatives of the order B, if point x is a member of the element B written as $x \,\epsilon\, B$. The first part of the equation is the polynomial approximation which we seek in this research while the other part known as the remainder is the approximation error. The quantity $(x - x_0)$ is the change in the variable and represents a deviation from x_0. If the Taylor's series expansion is to infinity the remainder becomes zero and the generic form of the expansion becomes:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^n(x_0)}{n!} (x - x_0)^n \qquad (4.6)$$

For this research the deviations are $\triangle m$ and $\triangle \theta$. We will truncate our Taylor's series at the first derivative because in economics, for simplicity, economic models are assumed to be linear, also because we assume accuracy of our approximation (Chiang and Wainwright 2005; Myers 1988 cites Newbery and Stiglitz, 1981).

Following Myers (1988), and, Holloway and Ehui (2001) and for ease of understanding we can rewrite the indirect utility function equation (4.4) in a more general sense thus:

$$V \equiv V(\theta,\, p,\, m) \qquad (4.7)$$

Hence, a Taylor's series expansion of equation (4.7) about the optimum point gives:

$$V(\cdot) \equiv V(\theta^*, p^*, m^*) = V_\theta \triangle \theta + V_p \triangle p + V_m \triangle m = 0 \qquad (4.8)$$

The changes in malaria prevalence, prices, and, the household endowmment income (which is a measure of willingness to pay for malaria abatement) are denoted by $\triangle \theta$, $\triangle p$, and, $\triangle m$ respectively, while, V_θ, V_p, and V_m are the first order Taylor series.

Since price, p, is constant thus $\triangle p = 0$ and equation (4.8) becomes:

$$V_m \triangle m = -V_\theta \triangle \theta \qquad (4.9)$$

The first order expressions for V_θ and V_m are arrived at from a simple application of the envelope theorem to the maximum-value function, equation (4.4), above thus:

$$V_\theta = -\lambda^* p_a f(L,\ A) \theta_j^{*'} \tag{4.10}$$

$$V_m = \lambda^* \tag{4.11}$$

The quantity $\theta_j^{*'}$ is the measure of inefficiency of the household as a result of malaria.

We can then estimate Δm which is a measure of the amount the household is willing to pay for malaria abatement from equation (4.9) resulting in the *Roy's identity* which is given as:

$$\Delta m = -\frac{V_\theta'}{V_m'} \Delta\theta \tag{4.12}$$

written in more elaborate form as:

$$\triangle m = \frac{(-) - \lambda^* p_a f(L,\ A)\theta_j^{*'}}{\lambda^*} \triangle\theta$$

The quantity $\Delta\theta$ is the malaria prevalence in the area of choice (which serves as a proxy for the possession of any prophylactic measure by the household) the value of which is chosen by the researcher.

We have all the ingredients to arrive at a reliable estimate that the household is willing to pay for malaria abatement. This amount is given as:

$$\Delta m = p_a . f(L,A).\theta_j^{*'} .\Delta\theta \tag{4.13}$$

The quantity Δm in equation (4.13) is the willingness to pay for malaria abatement. Its calculation will be taken up more elaborately in Chapter 8 but it is enough to say that it is the product of the price of home produced goods, the production technology using just land and labour, the inefficiency of the household with health as a major input in its calculation and the malaria prevalence (this is a constant value fixed by the researcher).

From equation (4.13) it can be seen that the quantities on the right can readily be arrived at by the application of our data and empirical estimation. Also, equation (4.4) presents

to us a direct relationship between the quantity on the right and the quantity on the left which can be readily seen by the reader. The empirical estimation of this is what we will take up in Chapter 8. However, what we have shown the reader is that a reliable estimate of the willingness to pay for malaria abatement is readily available by the simple application of the household model of Singh et al.(1986 pages $17 - 19$), the Taylor's series, envelope theorem and the Roy's identity. The reader will agree that this is simpler and more accurate than the use of the contingent valuation technique. Also, it is noteworthy to say that this technique still receives little attention in the areas of applied and development economics.

4.3. Summary

In this chapter, we have been able to motivate our conceptual framework by modifying Singh et al. (1986) work explained in chapter 3. From equation (4.13), the reader would agree that it is a direct method and an easy way to arrive at the willingness-to-pay value as compared to the Contigent Valuation technique. The Contingent Valuation technique is based on asking the farmers how much they are willing to pay for malaria abatement, which is subjective in nature, while, our method calculates this quantity directly which makes our approach objective.

We reiterate that the malaria variable can either be introduced into the frontier, inefficiency part of the model or on both sides.

Following from the epidemiology of malaria in chapter 2 above, the number of hours of labour time spent on the farm does not say the farmer will/will not be infected/affected by the malaria parasite. This is reflected in the way time was introduced into equation (4.1) above and explains the reason why the time variable 'disappears' in our comparative statics analysis above. We want to emphasize that our presentation is basic and it is intentional, with the sole aim of driving home our point, this is because, the introduction of so many variables (for example, in the production and inefficiency part of equation 4.1) may tend to deviate the attention of the reader from the main focus of our remit in this research (In chapter 8, we expatiate on our model to include other variables). We recapitulate that the core of our work is specifically on malaria and how it affects farmers' efficiency and not on 'general health'.

Our method involves the use of a very important technique in development economics called the stochastic frontier method, otherwise called the composed-error method. Thus, in the next chapter, we peruse the literature on the stochastic frontier analysis and further explain its importance to our research.

5. The Stochastic Frontier Model (The Composed Error Model)

In the last chapter we presented our conceptual framework, where, we emphasized the importance of the composed error model. Therefore, in this chapter, to further appreciate the essence of this model to our research, we continue our perusal of the literature with focus on the stochastic frontier model, otherwise called, the composed-error model. To reiterate, the two models - the household and the composed error models - are the fulcrum upon which equation (4.13) in chapter 4 rests.

Thus, this chapter scrutinises the literature on the stochastic frontier model, emphasize the thematic developments of the model over the years, and, espouse its use in understanding the rural household productive efficiency. Overall, we aim to emphasize the importance of the stochastic frontier model and to state where our conceptual framework links with extant literature on measuring productivity and efficiency (which the second part of equation (4.13) above focuses on).

5.1. Thematic Issues

Authors have made effort to develop the theoretical underpinning on the concept of productive efficiency (see for example; Leibenstein 1966; Comanor and Leibenstein 1969; Roll 1985). Also, it generates arguments between researchers, for example, Stigler (1976) and Leibenstein (1978), and, Schwartzman (1973) and Leibenstein (1973) show the theoretical controversy associated with the concept.

On the other hand, there has been a lot of success in the empirical application of the concept in the literature. An early impetus into the application of the concept of productive efficiency is the seminal work of Farrell (1957). Koopmans(1951)'s "activity analysis" and Debreu(1951)'s "coefficient of resource utilisation" inspire Farrell. Farrell (1957) dichotomises total efficiency into two; based on obtaining the maximum output from a set of input mix (technical efficiency), and, on the ability of the firm to choose the best set of input mix considering their prices, which he refers to as price (allocative) efficiency. In reality, decisions on these two indices are taken jointly by the firm *ex ante*, hence, when decisions on technical efficiency are taken, decisions on allocative efficiency are taken

simultaneously and *vice-versa* (Kopp, 1981). We present Farrell's model of efficiency measurement in figure 5.1 below:

Figure 5.1.:   Farrell's Model of Efficiency Measurement
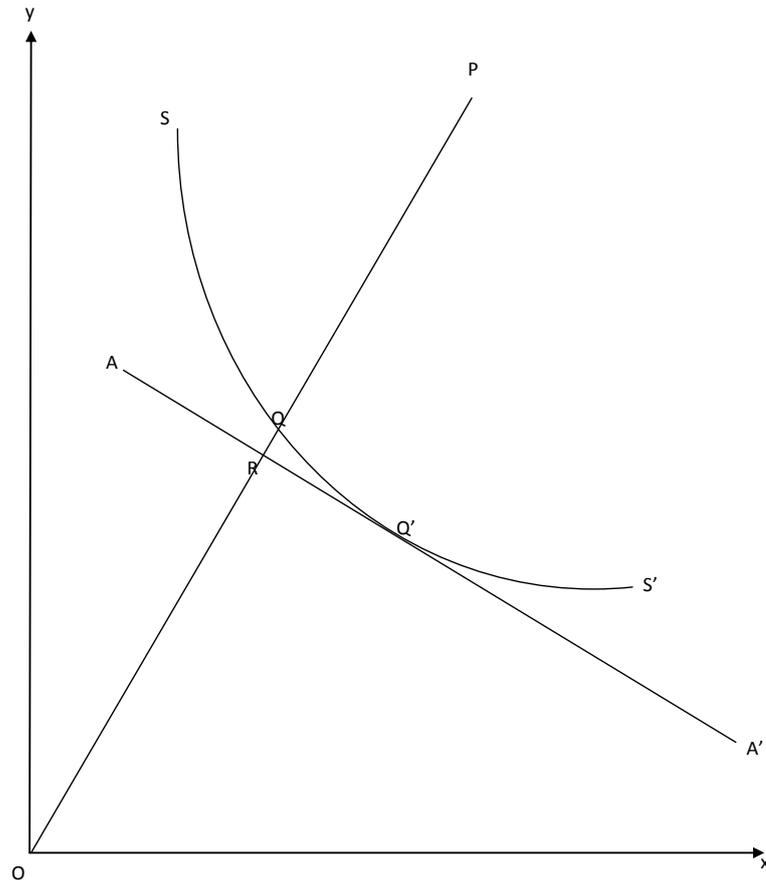
Initially, he considers a two input firm producing one product and operating at constant returns to scale of production. Assuming the efficient production function is known, the

isoquant SS' represents the various combinations of the two inputs that the perfectly efficient firm uses to produce unit output. In other words, it is the production function upon which the best combination of inputs is achieved. Thus, any point, such as Q, along the isoquant defines a technically efficient firm while any point above the isoquant, say point P, defines a technically inefficient firm. The distance, QP, along the ray OP represents the technical inefficiency of the firm at point P. In other words, the firm at point Q uses as much as OQ/OP amount of each input to produce the same amount of output as the firm at point P. In other words, firm Q, produces OP/OQ times more output that firm P. The point OQ/OP is thus the technical efficiency of the reference firm, that is to say, $1-OQ/OP$, which equals the ratio QP/OP is the technical inefficiency of the firm at point P.

It is seen from figure 5.1 that price of family labour is given, which in practice is not the case. The price of labour in the household is equal to the marginal value product of family labour. However, supposing information on prices are available and the firm achieves a perfect cost minimisation as one of its behavioural objectives, then, AA' represents the budget (iso-cost) line whose slope is the ratio of the prices of each input and the point of tangency, Q', is the least cost combination point (also called the allocative efficiency point). If the reference firm keeps its technical efficiency constant, and it is able to move from the not allocatively (price) inefficient point, Q, to the allocatively efficient point, Q' without any market barrier, then it will be able to reduce its cost by the ratio OR/OQ, *ceteris paribus.* At point P, the allocative inefficiency value for the firm at point P is given as RQ/OR.

Farrell (1957) refers to the point at which a firm is both technically and allocatively (price) efficient as *overall* efficiency, which is now referred to in the literature as *economic* efficiency. He states that it is the product of the technical and allocative efficiency. Hence, for our reference firm, its *economic* (overall) efficiency is:

$$E \bullet E = T.E \times A.E = {}^{OQ}\!/_{OP} \times {}^{OR}\!/_{OQ} = {}^{OR}\!/_{OP} \tag{5.1}$$

where *E.E, T.E*, and *A.E* are economic, technical, and, allocative efficiencies respectively.

The fraction, OR/OP, is defined as the ratio by which the reference firm's cost will reduce, if it were both technically and allocatively efficient.

Farrell(1957)'s study could not explain the situation of a multi-output firm, as well as, the case of an increasing or decreasing returns to scale, which he says, there is "no entirely satisfactory way of allowing for diseconomies of scale" (Farrell, 1957 p. 258). He notes that the most significant contribution of the paper is the decomposition of efficiency

into technical and allocative efficiency. The above description of efficiency by Farrell is an input-based analysis under strong regularity conditions, (he also presents the output measure of technical efficiency in this study).

Färe, Grosskopf and Lovell (1985 and 1994) present a more elaborate output measure of technical efficiency up to the multiproduct case including the assumption of weak regularity conditions. Fãre and Lovell (1978) state that the output and input measures of technical efficiency are equal under the constant returns to scale assumption.

However, Fãrsund and Hjalmarsson (1979) and Kopp (1981) state that this is not the case under non-homothetic production functions (non-constant returns to scale). Fãre, Grosskopf, and, Lovell (1985 and 1994) present an output-oriented approach to the measurement of allocative efficiency, also, Lovell (1993) and Fried and Lovell (2008) present a measure of allocative efficiency from a revenue efficiency perspective. Kumbhakar (1987 and 1989) analyse allocative efficiency under the profit maximisation framework, while, Schmidt and Knox Lovell (1979) and Kumbhakar and Tsionas (2005) examine allocative efficiency under the cost minimisation framework.

Kopp (1981) states three major features of Farrell (1957)'s work that have defined the literature on efficiency up to this present time. First, Farrell (1957) isoquant focuses on the primal production function which relates several inputs to output. The presentation of its dual, the cost function, would have been difficult using his approach considering the fact that the main objective was to dichotomise efficiency into technical and allocative efficiency. The cost function will serve as the derivation of technical efficiency, however, derivation at the allocative efficiency will have been precluded. Second, Farrell's isoquant is dependent on a set of production organisations using the same technology and producing the highest output from several inputs. The efficiency of a firm is calculated by comparing its performance with those in the set of production organisations. Third, the utilisation of a unit isoquant apart from being linearly homogeneous, it is also deterministic.

One question that arose in the early literature on productive efficiency is how much of input and output should be included in the analysis of productive efficiency? Lovell (1993) cites Knight (1933) who argues that this should not be the question as; if all inputs and outputs are included in an efficiency analysis, then the question of inefficiency should not arise because all production units will then have same productivity score. He redefines productivity to be the ratio of effective output to input.

We have presented how Farrell's (1957) work has influenced the study of efficiency over the years. The use of distance function in efficiency analysis has also developed over the years (see Cornes 1992 and Rodriguez-Alvarez and Knox Lovell 2004 for an elaborate exposition on this subject).

Another influential paper that has motivated research on efficiency in developing countries is Schultz (1964). His proposition that farmers in developing countries are "poor but

efficient" ignites a lot of study into the confirmation of this proposition. Welsch (1965), Chennareddy (1967), Lipton (1968), Adams (1986), Sauer and Mendoza-Esclante (2007) attempt to confirm this assertion (see Ball and Pounder 1996 for a critical analysis of this proposition). Because this proposition is embroiled in a lot of argumentation, we digress to review the literature in this area.

Chennareddy (1967) utilizes the linear regression analysis on a household data of one hundred and four rice and tobacco farmers in Southern part of India using a Cobb-Douglas production function. His findings are in accord with Schultz's hypothesis. He recommends that South Indian farmers should adopt modern technology and extension education in order to move to a higher frontier. Lipton (1968) disagrees with this recommendation. He argues that, if Schultz's findings are correct, then the rural farmers do not need any expert advice to improve their productivity, in other words, moving to a higher frontier should not be a problem for them. He further queries Schultz's assertion believing that it only works under a neo-classical theory of perfect competition; he affirms that if Schultz uses linear programming to analyse his data his findings will show that the rural farmer is inefficient. Welsch (1965) in his study on Abakaliki rice in eastern Nigeria makes use of the linear regression to affirm that peasant farmers respond to economic inducement by allocating efficiently among the several resources at their disposal, hence, he supports Schultz's hypothesis.

Sauer and Mendoza-Esclante (2007) seek to reconcile these diametrically opposing schools of thought. It puts to use a parametric normalized generalized leontief (GL) profit function technique to analyse joint production of Cassava flour and maize by small-scale farmers in Brazil. The small-scale farmers are allocatively efficient, they assert, but they show considerable inefficiency in the scale of operation. Schultz (1964)'s assertion not only instigates research in development and resource economics, but it also prompts research in anthropology and sociology (see Adams, 1986 and the review by Michelena, 1965 pp. 540-541).

The reader should note that the linear regressions of Chennareddy (1967) and Welsch (1965) give the shape of the technology of an average farm in the industry while the stochastic frontier model gives the shape of the best practice existing technology in the industry against which the efficiency of every other farms are measured relative to (Coelli 1995); in other words, Chennareddy (1967) and Welsch (1965) use an average response model for their analysis.

The technique of measurement of efficiency seems to be the more dominant topic in the literature than the theory. There are different techniques by which the measurement of efficiency has been classified in the literature. One classification is based on whether a functional form is specified *a priori* or not, in which case, efficiency measurement could be parametric or non-parametric. In the parametric technique, the researcher imposes a functional form on the frontier. The parametric technique is further divided into determ-

inistic and frontier methods. With the non-parametric, the frontier is calculated from the dataset, in the words of Koop (2003), the researcher is "letting the data speak". This, he says is very difficult to achieve as even in the non-parametric system, just like in the parametric, one need to impose certain structure on a particular problem in order to achieve ones objective(s). The use of the Data Envelopment Analysis (DEA) (another technique is the Free Disposal Hull, FDH) overshadows every other technique in the non-parametric class.

Charnes et al. (1978) introduce this technique and give the name as we know it today. The data envelopment analysis technique uses the mathematical programming technique to generate a piecewise envelope over the data points. The technique is widely used in technical efficiency studies, but it has the shortcoming of not incorporating randomness in measuring efficiency. As a result, it is not a statistical technique and therefore statistical inference cannot be carried out on it. Moreover, the envelope curve is not everywhere differentiable, which means a unique maximum point can not be arrived at.

Another method of classifying efficiency in the literature is according to the tool used to analyse them. The tool of analysis used could be mathematical programming or econometric in nature. While deterministic frontier models are analysed either by using the mathematical programming or the econometric techniques; the stochastic frontier models are only analysed by econometric techniques. Our focus in this research is the parametric technique of measuring technical efficiency, hence, we provide a critical review of the literature in this area.

5.2. The Parametric Frontier Method

5.2.1. The cross-sectional framework

The presentation of the theoretical derivations associated with the parametric efficiency measure is *sine qua non* to a better understanding of the *modus operandi* of this method by the reader. Consequently, in this section, we present the deterministic and stochastic frontier production functions and the relevant theoretical derivations of the measure of efficiency associated with each one of them. First, we present the annals of the deterministic frontier production function, immediately afterwards, we chronicle the stochastic frontier model.

5.2.1.1. The deterministic frontier production function

In table 5.1 below, we present, in chronological order, influential papers and different estimation techniques that define efficiency measurement since Farrell(1957)'s work. We

detail the type of functions utilized, the assumptions made as regards the disturbance term, the side constraints, the estimation methods, efficiency measure and data analysed. Farrell(1957)'s sole aim was to evaluate the performance of a firm. However, the frontier production function provides information on the prevailing best practice technology, as well as, serving as a measure of efficiency (Kopp 1981).

The primal production function of a parametric frontier is given as:

$$Q_i = h(X_i, \beta).TE_i \tag{5.2}$$

where $i = 1...N$ firms, Q is the single output, X is a vector of observations on n inputs, β is a vector of unknown production parameter, $h(.)$ is the production frontier, TE_i is the output oriented measure of technical efficiency of firm i which is arrived at by making TE_i the subject of equation (5.2) and defined as the ratio of the observed output to the maximum feasible output at the prevailing industry technology,

$$TE_i = \frac{Q_i}{h(X_i, \beta)} \tag{5.3}$$

$$TE_i \leq 1$$

Equation (5.3) shows the deviation of output from the frontier which serves as a measure of technical efficiency. This differs from the Farrell's measure where output is held constant and input is used as a measure of technical efficiency. We re-specify equation (5.2) based on technical inefficiency to give:

$$q_i = h(X_i, \beta). \exp(-u_i) \tag{5.4}$$

$$u_i \geq 0$$

u_i is a measure of technical inefficiency. The imposition of the restriction $u_i \geq 0$ ensures $TE_i \leq 1$. In other words, $TE_i = 1\text{-}u_i$.

Equation (5.2) has two basic characteristics: (I) It assumes all disturbances to be as a result of technical efficiency, these disturbances are either stipulated as unspecified random disturbances or follow a one-sided distribution like the truncated normal, gamma distributions (II) It is 'deterministic'.

The literature utilizes either the Mathematical programming or econometric techniques to obtain estimates for the production function specified in equation (5.2 to 5.4). The first set of literature in this area uses the mathematical programming technique to obtain estimates for the parameters (Aigner and Chu 1968; Seitz 1970; Timmer 1971). More recent applications of mathematical programming technique include Charnes et al. (1978), Fãrsund and Hjalmarsson (1979), Banker et al. (1984); Byrnes et al. (1984); Bjurek et al. (1990).

The mathematical programming technique involves the construction of a non-parametric piecewise surface to envelope the data and efficiency is measured relative to this surface. This method does not require functional and distributional assumptions for the efficiency term. The main disadvantage of this approach is that it does not lend itself to statistical testing. This is because the parameters are calculated and not statistically estimated; consequently, statistical inference and hypothesis testing are inhibited. This disadvantage is corrected for in the econometric method of efficiency analysis. Varian (1985) attempts to account for the measurement errors and introduce statistical testing into the non-parametric methods. His method involves using the minimum disturbances in the data that satisfies the inequalities stated in the production economics theory. He states that his method can be used in stochastic specifications.

Land et al. (1993), and, Olesen and Petersen (1995) introduce a type of mathematical programming that accounts for random noise in the data without them being stochastic called the "chance-constrained data envelopment analysis". However, Kalirajan and Shand (1999) states that its dependence on 'heavy' data, and, over-reliance on information on accurate data, and, the willingness of firms to venture into risks are its major disadvantage. Also, they state that some statistical inferences cannot be made from this technique.

The econometric methods estimate the parameters statistically and statistical inferences are made from these estimates. There are two main types of econometric techniques; these are the deterministic and stochastic econometric techniques. The deterministic econometric technique like the mathematical programming generates a hull from the subset of the sample over the entire data. The hull generated serves as the maximum best practice for the industry. Unlike, mathematical programming it estimates rather than calculate inefficiencies for the farms in the industry. This is because it makes distributional assumptions on the parameters, in order to, arrive at estimates. Thus, it enables statistical inferences to be made from the deterministic frontier. However, it still has the same disadvantage as the mathematical programming technique because it uses some and not all the data in the sample to generate the production frontier (hull). Also, it assumes all deviation from the hull (production frontier) is due to inefficiency without accounting for measurement and observation errors.

Also, as stated earlier, the stochastic econometric techniques assumes inefficiency is due to a one sided random error term. Different econometric methods exist in the literat-

ure in estimating the deterministic frontier. These include the Modified Least Squares (MOLS) of Richmond (1974), Corrected Least Squares (COLS) of Winsten (1957), and the maximum likelihood estimator (MLE). Next, we discuss the practical application of the deterministic frontier.

5.2.1.2. Practical Application of the Deterministic Frontier

The deterministic frontier has been applied in the literature. One of the early applications of the deterministic frontier are Shapiro and Mãller (1977), Shapiro (1983), and, Belbase and Grabowski (1985). Shapiro and Mãller (1977) attempt to estimate the technical efficiency of forty farms in Geita district of Tanzania. They follow Timmer(1971)'s method of analysing technical efficiency by applying linear programming to a Cobb – Douglas production frontier. Their result which is similar to that of Chennareddy (1967) shows that the traditional farmer can improve his technical efficiency by adopting modern farming practices through easy access to information.

This, they say, will be at the expense of non-economic costs like the farmer being branded "unsociable" by his community. Shapiro (1983) working in the same district as Shapiro and Mãller (1977) tries to confirm the 'poor, but efficient' hypothesis of Schultz (1964) but discovers that the hypothesis may not be applicable in terms of peasant agriculture in Tanzania because their output could still be increased. They remark that Schultz's hypothesis will hold if every other farmer has the same efficiency as the most efficient farmer in the sample. These assertions echo the conclusion of Lipton (1968). He uses the same model and method of analyses as Shapiro and Mãller (1977).

Belbase and Grabowski (1985) utilize the Corrected Ordinary Least Square (COLS) approach of Winsten (1957) on a cross-sectional sample of farms in Nuwakot district of Nepal. They record an average technical efficiency value of eighty percent for joint production of rice, maize, millet, and, wheat. The average technical efficiency value of individual frontier calculation for rice and maize is given as eighty-four percent and sixty-seven percent in that order. They find correlation between technical efficiency and other variables which are nutritional level, income, and, education. Technical efficiency is however not correlated with farming experience.

Some studies utilize the deterministic model to investigate the impact of certain agricultural policies on productivity. One expects these policies to actually increase productivity *a priori*, but this is not always the case. One of such studies is Taylor et al. (1986). They employ a deterministic production function and discover that the World Bank sponsored credit programme - PRODEMATA - did not impact positively on the technical efficiency of farmers in *Minas Gerais*, Brazil. Their result shows that there is no difference between the technical efficiency of farmers who participate in the programme and those who did

not participate. This paper is one of the few that compares both the results of the Corrected Ordinary Least Square and the maximum likelihood approaches. Unexpectedly, the participant farmers in the PRODEMATA programmes have a slightly lesser allocative efficiency than non-participant farmers supporting Schultz's hypothesis. Amara et al. (1999) use the deterministic frontier to discover the relationship between technical efficiency and the adoption of conservation technologies by potato farmers in Quebec. They find that farming experience and the adoption of conservation technologies have a positive influence on technical efficiency.

Wouterse (2010) uses variable returns to scale data envelopment analysis of the type of Seitz (1970) to explore the impact of migration on cereal farmers' technical efficiency in Burkina Faso . She observes that intercontinental migration improves remittances to productive capital of cereal farmers. On the other hand, it prevents them from moving to a higher frontier because they cannot expand land for farming. She remarks that remittances will not be enough to improve agriculture in Burkina Faso; and recommends that policy makers need to investigate the imperfections in the labour market of the migrant-sending economies.

Aly et al. (1987) assume a ray-homothetic function to analyse a sample of eighty-eight Illinois grain farmers using the corrected ordinary least square approach. They estimate a deterministic statistical frontier. They observe that the farms produce up to fifty-eight percent of their potential capabilities. The farms have mean technical inefficiency values of sixty percent and scale inefficiency values of forty percent. They also observe that the technical efficiency values are positively correlated with the size of the farm. In other words, large farms are more efficient than small farms.

So far, we have discussed the parametric frontier model, however, because the deterministic frontier model takes all deviation from the frontier has inefficiencies, it over-values inefficiency estimates. Hence, it can be argued that, it does not provide an accurate measure of efficiency, for example, Taylor and Scott Shonkwiler (1986) discover that the deterministic frontier gives over seventy percent inefficiency values while the stochastic frontier gives twenty percent value for inefficiency. The advent of the stochastic frontier models corrects for this major drawback which we will discuss in the next section.

5.3. The Stochastic Frontier Models

Aigner et al. (1977), Meeusen and van den Broeck (1977), and, Battese and Corra (1977) simultaneously introduce the stochastic frontier analysis as we know it today. Apart from incorporating the efficiency term into the deterministic model, it also includes the effect of measurement errors, hence, the name 'stochastic'.

The primal production function of a parametric frontier is given as:

$$Q_i = h(X_i, \beta).\epsilon_i; \ \epsilon_i = v_i - u_i \tag{5.5}$$

The model is also referred to as the composed error model because of the presence of the $(v_i - u_i)$ term. The term, v_i, is the randomness (or statistical noise) associated with the data, while, u_i, is the measure of technical inefficiency. It is common practice in the literature to assume v_i to be normal, independent and identically distributed. With reference to u_i, it is one-sided and the literature proposes several distributions for it. The distributions proposed include the half-normal of Aigner et al. (1977), the exponential of Meeusen and van den Broeck (1977), the truncated normal of Stevenson (1980), and, the gamma distributions of Greene (1990).

Classical econometricians often use half-normal, exponential, Gamma, and, Truncated - Normal distributions, while, the Bayesian econometricians are more akin to using the Truncated - Normal and Gamma distributions. Other things being equal, the literature assumes that the two types of error terms are independent of each other. This necessitates the need to find the expected conditional form of our choice parameter, u_i given the composed error variables, $v_i - u_i$. Jondrow et al. (1982) and Kalirajan and Flinn (1983) independently derive the conditional form of u_i for the half-normal. The Jondrow et al. (1982) version is given as:

$$E[u_i|\epsilon_i] = \sigma_* \left[\frac{h(\epsilon\lambda/\sigma)}{1 - H(\epsilon\lambda/\sigma)} - \frac{\epsilon\lambda}{\sigma} \right] \tag{5.6}$$

where $h(.)$ is the density of the standard normal distribution, $H(.)$ is the cumulative density function, $\lambda = \sigma_u/\sigma_v$; $\epsilon_i = v_i - u_i$, $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\sigma_* = \left(\sigma_u^2\sigma_v^2/\sigma^2\right)^{1/2}$.

This said, one can compute the technical efficiency from the conditional estimates for each firm as:

$$TE_i = 1 - E\left[u_i|\epsilon_i\right] \tag{5.7}$$

Battese and Coelli (1988, p.390) state a preferable way of writing equation (5.6) as $\exp\{(-u_i)|e_i\}$, this is because equation (5.6) is a first order approximation of the infinite series $\left(\exp\{(-u_i)|e_i\} = 1 - u_i + u_i^2/2 - u_i^3/3!...\right)$ and the remaining part might be significant if u_i is not close to zero. They propose an alternative mean efficiency estimator expressed as:

$$E\left[\exp(-u_i)|e_i\right] = \frac{1 - \Phi\left[\sigma - (\mu/\sigma)\right]}{1 - \Phi\left(-\mu/\sigma\right)} \cdot \exp(-\mu + \frac{1}{2}\sigma^2) \tag{5.8}$$

As noted in Jondrow et al. (1982, p. 235), intrinsic variability independent of sample size exists due to the conditional distribution of u given ϵ_i and this is because ϵ_i contains imperfect information about u. As a result, Waldman (1984) states that the estimators of technical efficiencies (along with other estimators) obtained by Jondrow et al. (1982), Kalirajan and Flinn (1983), and, [Battese and Coelli (1988)] are not consistent. For the sake of further inferences, confidence intervals for each of the point estimates above can be carried out. Hjalmarsson et al. (1996) arrive at confidence intervals for the Jondrow et al. (1982) technical efficiency estimator, while, Bera and Sharma (1999) obtain the confidence interval for the Battese and Coelli (1988) estimator.

Horrace and Schmidt (1996) derive confidence intervals for different levels of uncertainties for the Jondrow et al. (1982) and the Battese and Coelli (1988) models under the panel data framework, while, Jensen (2000) carries out similar analysis using the Corrected Ordinary Least Squares (COLS) approach. Horrace and Schmidt (1996) develop the "multiple comparisons with the best" methodology and advance this procedure in Horrace and Schmidt (2000).

Jondrow et al. (1982) also compute the moment for u_i conditional on the composed error term for the exponential case. This is given as:

$$E[u_i|\epsilon_i] = \sigma_v \left[\frac{h(\epsilon_i/\sigma_v + \sigma_v/\sigma_u)}{1 - H(\epsilon_i/\sigma_v + \sigma_v/\sigma_u)} - (\epsilon_i/\sigma_v + \sigma_v/\sigma_u)\right] \tag{5.9}$$

Because the half-normal and exponential distributions have a mode at zero, as a result, Murillo-Zamorano (2004) states that they give conditionally high efficiency levels, especially for efficiency scores about the zero point. Moreover, he asserts, that the shape of the distribution is predetermined, which he considers as a major disadvantage.

Stevenson (1980) provides a truncated-normal scenario, which is a generalisation of the half-normal case. Also, Greene (1990) extends Greene (1980) proposition for the deterministic Gamma frontier model to the stochastic frontier context. He presents a stochastic composed error with the gamma density of the form:

$$Q_i = h(X_i, \beta_i) + v - u, \tag{5.10}$$

where; $v \sim N(0, \sigma^2)$ and $u \sim G(\Theta, P)$;

Θ, $P > 0$, $u \geq 0$, Θ is a scale parameter

He computes the conditional moment of u_i as:

$$h(u|\epsilon) = \frac{\left(2\pi\sigma^2\right)^{-1/2}\left(\exp(-1/2\left[u + (\epsilon + \Theta\sigma^2)\right]^2/\sigma^2)\right)u^{P-1}}{Prob\left[\mathbb{Q} > 0|\epsilon\right]h(P-1,\epsilon)} \tag{5.11}$$

He also shows that the truncated normal distribution for the conditional moment of u is a special case of equation (5.9) when u has an exponential distribution (see Greene 1990, p. 147 for further explanation). Murillo-Zamorano (2004) states that the gamma distribution is not as popular as the normal-half-normal and the exponential distributions in maximum likelihood inference because it is complex and involves a lot of computational drudgery. Moreover, the shape parameter is hard to estimate in small sample sizes which greatly affects the stochastic frontier estimates, the individual efficiency values, and, the assignment of the overall variance to the stochastic frontier and inefficiency values (Ritter and Simar 1997). On the contrary, the Normal-Gamma and the Normal-Truncated-Normal cases are the model of choice in Bayesian econometrics.

The stochastic frontier functions can also be analysed using the approach of the method of moments. Several authors present the stochastic frontier functions under varying distributional assumptions for this approach. Olson et al.(1980, p.80-82) for the normal-half-normal case, Harris (1989) presents the normal-truncated-normal context (Murillo-Zamorano 2004 cites Harris, 1992) while Greene (1993, 1997, and, 2008) present the Normal-Exponential and Normal-Gamma context respectively. We depict all of these developments in table 5.1 below:

Table 5.1.: Historical Developments that have Shaped Frontier and Effiiciency Measurement in the Early Literature

Study	Production Function	Form of disturbance	Side constraints	Estimation method	Efficiency measure	Data type
Aigner and Chu (1968)	Homothetic Cobb-Douglas:$q_i = AX_1^{\hat{\beta_1}} X_2^{\hat{\beta_2}} u_i$	efficiency subsumed in random shock, u *(deterministic)*	$AX_1^{\hat{\beta_1}} X_2^{\hat{\beta_2}} \geq q_i$	mathematical programming	None specified	single output-cross-sectional data
Seitz (1970)	Non-homothetic: Maximize $u = \Sigma_i^n U_i$	efficiency subsumed in random shock, u *(deterministic)*	$\beta_i \geq 0$ and $\sum_i^n \widehat{\beta_i} X_i \leq q_i$	Linear programming	None specified	Single output
Timmer (1971)	Homothetic: minimize $u = \Sigma_{i=1}^n U_{it}$	efficiency subsumed in random shock. *(deterministic)*	$\beta_i \geq 0$ and $\sum_{i=0}^n \beta_i X_{jit} \geq Q_{it}$	Linear programming	$\frac{Q_{it}}{\hat{Q}_{it}}$ =index of efficiency for each firm. Average efficiency over time = $\frac{1}{n} \sum_{i=1}^n \frac{Q_{it}}{\hat{Q}_{it}}$	Aggreagted output - Panel data

(*Adapted and modified from Kopp, 1981*) **(continued)**

Study	Production Function	Form of Disturbance	Side Constraints	Estimation Methods	Efficiency measure	Data type
Afriat (1972)	Homothetic Cobb-Douglas: $q_i = h(X_i, \beta).e^{-u_i}$ *(deterministic)*	$u_i \sim G(u_i, n)$	None	suggests Maximum Likelihood Estimation	mean sample efficiency, $E(exp(-u_i)) = \int_0^\infty e^{-u}.G(u_i,n)du$	None stated
Richmond (1974)	Homothetic Cobb-Douglas:$q_i = h(X_i, \beta_i).e^{-u_i}$ *(deterministic)*	gamma distribition, $G(u_i, n)$	None	Adjusts intercept of OLS	mean sample efficiency	cross-sec- tional
Aigner, Amemiya, Poirier (1976)	Linear regression: $q_i = h(X_i, \beta_i) + u_i$ *(deterministic)*	$u_i = \begin{cases} \frac{u_i^*}{\sqrt{1-\theta}} & if\ u_i^* > 0 \end{cases}$ $u_i = \begin{cases} \frac{u_i^*}{\sqrt{\theta}} & if\ u_i^* \leq 0 \end{cases}$	None	Maximum Likelihood	θ is the average OLS function, and it measures the degree of variability in output above or below the mid-point, zero, attributable to inefficiency	simulated data
Schmidt (1976)	Homothetic Cobb-Douglas: $q_i = h(X_i, \beta) - u_i$ *(deterministic)*	$u_i \sim N(\mu, \sigma^2)$	$\mu > 0$	OLS and Maximum Likelihood	$u_i^* = u_i - \mu$	None

(continued)

Study	Production Function	Form of Disturbance	Side Constraints	Estimation Methods	Efficiency Measure	Data type
Meeusen and van den Broeck (1977)	Homothetic Cobb-Douglas: $q_i = h(X_i, \beta_i).e^{-u_i}$ *(stochastic)*	$\epsilon_i = v + u$ composed error, where $v \sim N(0, \sigma^2)$, $u \sim exponential$	None	Maximum Likelihood	$E(e^{-u})$. Average sample efficiency is ranges between 0 and 1	cross-sectional
Aigner, Lovell, and Schmidt (1977)	Linear $q_i = h(X_i, \beta_i) + \epsilon_i$ *(stochastic)*	$\epsilon_i = v + u$ composed error, where; $v \sim N(0, \sigma^2), u \sim N^+$ *(half-normal)*	None	Maximum Likelihood	$\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \sigma_u/\sigma_v$ $\lambda \to \infty$; full frontier $\lambda \to 0$; average function. Sample based measure of relative disturbance variation	cross-sectional
Stevenson (1980)	$q_i = h(X_i, \beta) + \epsilon_i$ *(stochastic)*	$\epsilon_i = v + u$ composed error, where; $v \sim N(0, \sigma^2)$, $u \sim N^+$ (truncated - normal)	None	Maximum Likelihood	$\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \sigma_u/\sigma_v$ $\lambda \to \infty$; full frontier $\lambda \to 0$; average function. Sample based measure of relative disturbance variation	cross-sectional
Greene (1980)	$q_i = h(X_i, \beta) - u$ *(determinisic)*	$h(u) \sim G(\lambda, P)$ $\equiv$ $\frac{\lambda^P}{\Gamma} u^{P-1} e^{-\lambda u}$(gamma)	$u \geq 0, \lambda > 0$, $P > 2$	Maximum Likelihood	$E(u) = \lambda/P$ $E(e^{-u}) = 2^{-P}$	cross-sectional
Greene (1990)	$q_i = h(X_i, \beta) + v - u$ *(stochastic)*	$h(u) \sim G(\lambda, P)$ $\equiv$ $\frac{\lambda^P}{\Gamma} u^{P-1} e^{-\lambda u}$(gamma)	$u \geq 0, \lambda > 0$, $P > 2$	Maximum Likelihood and Methods-of-Moments	$E(u) = \lambda/P$ $E(e^{-u}) = 2^{-P}$	cross-sectional

Study	Production Function	Form of Disturbance	Side Constraints	Estimation Methods	Efficiency Measure	Data Type
Olson et. al.(1980)	Cobb-Douglas: $y = X\beta + \epsilon$	$\epsilon = v - u$, where $v \sim N\left(0, \sigma_v^2\right)$, $u \sim N^{half-normal}\left(0, \sigma_u^2\right)$	None	Maximum Likelihood, Corrected Least Squares, Newton-Raphson, and, Method of Moments	$\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \sigma_u/\sigma_v$	Monte Carlo Experiment
Greene (1993,1997,2008)	$y_i = \alpha + \beta' X_i + \epsilon_i$ (stochastic)	$\epsilon_i = v_i - u_i$ u_i is either exponential, gamma, truncated-normal, or, half-normal	$u_i \geqq 0$, and, v_i is unrestricted	Maximum Likelihood, Corrected Least Squares, Modified Least Squares, and Method of Moments	$\lambda = \sigma_u/\sigma_v$, $\sigma^2 = \sigma_u^2 + \sigma_v^2$	Panel Data

It is noteworthy to state that the success of cross-sectional stochastic frontier models is dependent on the specification of a distribution for the statistical noise and the inefficiency parameters. These parameters are very sensitive to alternative choice of distribution, also, the identification of these parameters is dependent on the strength of the distributional assumptions made. Furthermore, the estimates, though unbiased, are inconsistent. The estimates are inconsistent because cross-sectional measures of inefficiency use data collected at a single point in time, thus, these estimates cannot be confirmed to approach the true parameter since the sample size is given. Consistency is possible if as the sample size increases, the variance gets smaller and its values approach the true parameter. These necessitate the use of panel data in the stochastic frontier analysis. Under the panel data framework, the strong distributional assumptions on these parameters are relaxed.

Schmidt and Sickles (1984) discuss three advantages of using the panel data models over the cross-sectional data models (Kalirajan and Shand 1999p. 159, and, Murillo-Zamorano 2004 echo these advantages). First, they state that no specific distributional assumptions are essential in order to arrive at consistent inefficiency estimates. This is because all the relevant parameters can be arrived at by use of the fixed-effects and random-effects model approaches. Second, the panel data model relaxes the assumption that the exogenous variables (input variables inclusive) and the inefficiency terms are independent. They state that this is an important advantage of the panel data model over the cross-sectional model. Coelli et al. (2005 p.285) corroborate this point by stating that the imposition of a specific distributional assumption on the inefficiency term will preclude the benefits associated with the use of panel data. They state that the independence assumption placed on the inefficiency term is not realistic because an efficient firm is expected to be efficient from one period to another *ceteris paribus*, also, a less efficient firm is expected to be efficient over time.

Third, they state that obtaining consistent technical inefficiency estimates for each of the sample units in the cross-sectional data is not possible (Jondrow et al. 1982 and Kalirajan and Flinn 1983 state this clearly) but these estimates are consistently available in the panel data models. This is because the continual updating of information on a particular unit over time causes noise to be averaged in the overall residual. For us, the third point is the most important advantage of the panel data over the cross-sectional data, because the art of updating goes with the Bayesian approach which we utilize in this research.

The traditional method of estimating panel data, that is, the fixed effects and the random-effects are also employed in the analysis of a panel stochastic frontier model (see Hsiao 2003 for a thorough exposition to the traditional fixed and random effects models). Meanwhile, for a fixed effects model, we consider a set of firms with systematic inefficiency and the individual firm specific inefficiency, u, are constant over time. Thus, the production function is written as:

$$Q_{it} = \beta_0 + X_{it}\beta + v_{it} - u_i \tag{5.12}$$

where $i = 1, 2, ..., N$

$t = 1, 2, ..., T$

Equation (5.12) can be written as:

$$Q_{it} = \beta_{0i} + X'_{it}\beta + v_{it} \tag{5.13}$$

where $\beta_{0i} = \beta_0 - u_i$

First, we obtain the 'within-groups' transformation, estimation is then carried out using the ordinary least squares approach using the deviation from unit mean as follows:

$$Q_{it} - \bar{Q}_i = \beta'\left(X_{it} - \bar{X}_i\right) + v'_{it} \tag{5.14}$$

The individual technical efficiency measure is given as:

$$TE = \exp(\hat{u_i}) \tag{5.15}$$

where $\hat{u}_i = \hat{\beta}_0 - \hat{\beta}_{0i} \quad i = 1, 2, ..., N;$

$\hat{\beta}_{0i} = max\left(\hat{\beta}_{0i}\right)$

From equation (5.14), it is observed that the 'fixed effects' model allows for correlation between the independent variables and the inefficiency term. Moreover, it is easy and straight forward and gives consistent estimates. However, Kalirajan and Shand 1999 state that care should be taken in interpreting the result because there is the likelihood that the firm-specific effects may include characteristics that actually vary over the firms but are time-invariant. This is the consequence of expressing the initial model in the form of deviation from the mean.

Simar (1992) applies the fixed-effects model to data on railways and discovers that it provides poor estimation of the intercepts and the slope coefficients which results in unacceptable inefficiency estimates. This impediment causes the random-effects model to be the model of choice in most literature. Murillo-Zamorano(2004 p. 51) is more specific on the drawback of the fixed - effects model. He states that if time - invariant

variables (like institutional factors, and, location factors) are included in the frontier, the fixed-effects model will prevent the usage of these variables, especially when the initial model is expressed using the deviation from unit mean approach. This is because they remain constant and the difference between each successive variable and the unit mean gives zero.

The random-effects model corrects for this shortcoming in time-invariant analysis by assuming independence of the inefficiency term and the regressors, as a result, some time invariant regressors can be included in the model. Thus, we rewrite equation (5.13) as:

$$Q_{it} = (\beta_0 - E(u_i)) + X_{it}\beta + v_{it} - (u_i - E(u_i)) \tag{5.16}$$

where v_{it} and $u_i^* = (u_i - E(u_i))$ are zero mean

Equation (5.16) can be analysed using the two-step generalised least squares (GLS) method if the regressors are not correlated with the inefficiency terms, this obtains consistent and unbiased estimates. On the contrary, if a correlation exists between the inefficiency term and the regressors, Hausman and Taylor (1981) propose estimating the model under the instrumental variable framework, in order to obtain consistent and unbiased estimates.

The use of the instrumental variable framework allows some of the regressors to be correlated to the firm-specific individual efficiency effects variable. In concise terms, the random effects model allows time invariant variables to be included in the technology specification. On the other hand, the imposition of the assumption that the inefficiency terms u_i's have to be uncorrelated with the regressors serves as the major advantage of the fixed effects over the random effect. Also, an advantage the instrumental variable estimation has over the generalised least squares estimation is that it allows the firm-specific inefficiency effects to be correlated with the exogenous variables in the equation.

As stated earlier, one of the advantages of the panel data over the cross-sectional data is the lack of the imposition of a distribution on the stochastic frontier model, however, there are situations in-which these distributions are known *a priori,* then the panel data can be estimated using the maximum likelihood techniques. Building on this, Pitt and Lee (1981) uses the maximum likelihood techniques to derive the normal-half-normal balanced panel data form for Aigner et al. (1977) cross-sectional model. Kumbhakar (1987, p. 344) and Battese and Coelli (1992) derive the normal-truncated-normal form for the stochastic frontier model which is the general specification for the normal-half-normal balanced panel data model. Battese et al. (1989) took the work of Battese and Coelli (1988) a step forward by applying the maximum likelihood method of estimation to unbalanced panel data using the *Davidon-Fletcher-Powell* algorithm. They present both

the normal-half-normal and the normal-truncated-normal case for the inefficiency term, u_i.

According to Murillo-Zamorano (2004), researchers are at a cross-road when it comes to comparing and choosing the best approach among the three (fixed-effects, random-effects and maximum likelihood models) approaches because these methods present divergent properties and impose varying requirements on the data. However, he opines that the choice of one over the other depends on the situation and the structure of the intended analysis. Kumbhakar and Lovell (2000) after reviewing the results of the analysis of Gong and Sickles (1989), Gathon and Perelman (1992); Bauer et al. (1993), Bauer and Hancock (1993), (Ahmad and Bravo-Ureta 1996) conclude that the conflicting evidence from these literature affirms that the three approaches are likely to generate similar efficiency rankings, especially at the top and the bottom of the distribution - where managerial expertise is concentrated.

Heretofore, we have analysed panel data under the time-invariant model, Schmidt (1985, p.313) states that it is not only an attractive assumption but also a powerful one because it allows the researcher to resolve some problems associated with estimating frontier models. Hence, we allow inefficiency to vary over time and this is the subject of the next section.

5.3.1. Technical (In)Efficiency under Time Varying Model

Like in the time-invariant model, technical efficiency using the time variant model is analysed using the fixed-effects, random-effects and maximum likelihood methods. Sickles et al. (1986) first make effort in analysing a time variant efficiency model, but as stated in Kalirajan and Shand (1999) their system of equations is highly parameterized and they use the maximum likelihood estimation procedure. The high parameterization of their model is because the introduction of inefficiency into a system of profit-maximizing output - supply and input - demand equations and adopting a maximum - likelihood estimation procedure. Also, they attempt to measure allocative efficiency without due consideration of technical efficiency. Kumbhakar (1990) presents a model of time-variant efficiency and observes that the time-invariant case is a special case of the time-varying case which must be tested by appropriate statistical tools. He suggests a model with half-normal distribution and defines the firm-specific inefficiency effect as the product of an exponential function of time with two parameters and a time invariant random variable:

$$Q_{it} = \beta_0 + \Sigma\beta_j X_{ijt} + \epsilon_{it} \tag{5.17}$$

where $\epsilon_{it} = v_{it} + u_{it} = v_{it} + \gamma(t)u_i$;

$i = 1, 2, ..., N;$ $\qquad$ $t = 1, 2, ..., T$

$\gamma(t) = [1 + \exp(bt + ct^2)]^{-1}$

The exponential time function could be monotonically increasing (or decreasing) or concave (or convex). The firm-specific inefficiency becomes time invariant when $b = 0$. He analyses the model using a maximum likelihood estimation procedure.

Battese and Coelli (1992) present a truncated normal variant of the Kumbhakar (1990) model. They consider a time-varying model with unbalanced panel data stated as:

$$Q_{it} = h(X_{it}, \beta) \cdot \exp(v_{it} - u_{it}) \tag{5.18}$$

such that

$$u_{it} = \eta_{it} u_i = \{\exp[-\eta(t - T)]\} u_i, \quad t \in \Lambda(i); \quad i = 1, 2, \dots, N$$

where η is an unknown scalar parameter, $\Lambda(i)$ is the set of T_i time periods among the t periods involved for which the observation for the ith term is obtained. This model does not require data for all the time periods in question. However, the rigid parameterization of the exponential specification of the firm-specific effect $u_{it}^{'s}$ is a major drawback of this model. This means as Kalirajan and Shand (1999) state; that technical efficiency must either increase or decrease at a decreasing rate ($\eta > 0$), decrease at an increasing rate ($\eta < 0$), or remain constant ($\eta = 0$). This rigidity is a consequence of the single parameter specification of the exponential time function; they recommend a two-parameter specification.

Cornwell et al. (1990) propose a model for the time-varying firm-specific effects with a distributional assumption. Thus, we opine that they are the first to provide a panel data generalisation to the Schmidt and Sickles (1984) model for the time-varying firm-specific effects. Theirs is a 'proper' time-varying panel data model. They specify the model as:

$$Q_{it} = \beta_0 + X_{it}'\beta + v_{it} - u_i \tag{5.19}$$

Equation (5.19) can be further simplified and written as:

$$Q_{it} = \beta_i + X_{it}'\beta + v_{it} \tag{5.20}$$

where $\beta_i = \beta_0 - u_i, i = 1, 2, \dots, n$

In other words, Cornwell et al. (1990) model does not require any assumption about

density function, to arrive at firm-specific effects. They state the time-varying firm specific effect as a quadratic function of time as:

$$\beta_{it} = \theta_{i1} + \theta_{i2}t + \theta_{i3}t^2 \tag{5.21}$$

Equations (5.20) and (5.21) can now be estimated using either the fixed-effects or random effects method through the Generalized Least Square, or the Instrumental Variable framework.

Lee and Schmidt (1993) consider a special case of the Kumbhakar (1990) model where $\gamma(t)$ is represented by a set of dummies for time using a modified format of the Holtz-Eakin et al. [1988, eq. (2.2), p. 1373] equation, such that, the firm-specific inefficiency effect at a different time period is the product of a set of dummies time effects and the individual firm technical inefficiency. This is given as:

$$u_{it} = \gamma_t u_i \tag{5.22}$$

where γ_t are a set of dummies for individual time effects, u_i is the individual firm technical inefficiency, and, u_{it} is either a fixed or random firm-specific effect.

The above method of measuring the stochastic frontier model uses the Frequentist approach to efficiency measurement as against the Bayesian method of measurement - we will elaborate on the basic differences between these two methods in the next section. Also, in this research, the problem of choosing between the fixed and random effects model is circumvented by the fact that we are using the Bayesian method of analysis. The Bayesian method, (we utilize this method in this thesis) as we will review in section (5.4), does not depict any differences between the fixed and random effects models. Thus, we present the Bayesian approach to measuring technical efficiency.

5.4. The Bayesian Approach to Measuring Technical Efficiency

In order for the reader to understand our remit in this section, we present a basic definition of terms used in Bayesian Econometrics.

Prior probability, $p(\theta)$: This represents the level of knowledge the researcher has about the parameter before seeing the data. If the researcher has a vague knowledge of the distribution of the parameter, then he adopts a *non-informative prior* for the parameter.

On the contrary, if the researcher has specific knowledge of the parameter distribution, he adopts a *Proper or Informative prior* for the parameter distribution.

Likelihood function: This is the data generating density and it is used to modify the prior probability, $p(\theta)$, and it is denoted by $p(y|\theta)$, when it is viewed as a function of y, and, $l(\theta|y)$, when viewed as a function of θ.

Posterior probability: The updated knowledge obtained from the modification of the prior belief by the likelihood function gives the posterior density denoted by $p(\theta|y)$. From the application of the standard conditional probability equation, the relation between these densities is given as:

$$p(y, \theta) = p(y|\theta) \cdot p(\theta) = p(\theta|y)p(y) \qquad (5.23)$$

The posterior density is given as:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \qquad (5.24)$$

where $p(y)$ is the *marginal likelihood* which is also be written as $m(y)$ and it is found by integrating over the product of the likelihood function and the prior densities

$$m(y) = \int p(y|\theta)p(\theta)d\theta \qquad (5.25)$$

Leaving out the marginal likelihood, equation (5.24) can be written as:

$$p(\theta|y) \propto p(y|\theta)p(\theta) \qquad (5.26)$$

Thus, the posterior becomes a function of the likelihood and prior density.

5.4.1. Basics of Model Selection in Bayesian Econometrics

The above methods of measuring technical efficiency described in earlier sections employed the *sampling theory* (classical) framework, however, an alternative method exists in the Bayesian approach. From the above sections, the reader will observe that different models have been developed to surmount one or more weaknesses in the analysis of productive

efficiency. However, one common problem that is observable is the need to overcome certain restrictive parametric assumptions and then comes the question of the best model to choose among all the models?

The procedure considers all the models proposed in the literature and compares them using their *posterior model probabilities*. Because of computational issues associated with the direct use of the posterior model probabilities, the *posterior odds ratio* or the *Bayes factor* is used in selecting the most appropriate model that fits the data. Also, instead of choosing the best model that fits the data, *Bayesian model averaging* has been developed in the literature and this involves averaging over the proposed models with their individual posterior model probabilities as weights.

Hence, if M_i represents a set of different m models, the posterior model probability conditioned on the data, y, $p(M_i|y)$, is given by:

$$p(M_i|y) = \frac{p(y|M_i)p(M_i)}{p(y)} \tag{5.27}$$

where $p(M_i)$ is the *prior model probability* and it represents a measure of the researcher's belief in a particular model before seeing the data, $p(y|M_i)$ is the *marginal likelihood*.

The *marginal likelihood* also called the *model evidence* or *integrated likelihood* is given thus:

$$p(y|M_i) = \int p(y|\theta^i, M_i)p(\theta^i|M_i)d\theta^i \tag{5.28}$$

Equation (5.28) shows that the marginal likelihood depends on the likelihood function and the prior belief of the researcher. Since $p(y)$ in equation (5.27) is hard to calculate directly, the *posterior odds ratio*, PO, can be used in place of the *posterior model probability*. These two quantities are related, the only difference is that the posterior odds ratio is the ratio of two posterior model probabilities. One can use the posterior odds ratio to compare two models, i and j, we have:

$$PO_{ij} = \frac{p(M_i|y)}{p(M_j|y)} = \frac{p(y|M_i)p(M_i)}{p(y|M_j)p(M_j)} \tag{5.29}$$

In addition, the researcher might want to place equal prior weights on each model, in other words, $p(M_i) = p(M_j)$, then, the equation (5.29) results in a ratio of the marginal likelihood. This ratio is referred to as the *Bayes Factor*, *BF*, and the results also helps the researcher in selecting the best model. Thus;

$$BF_{ij} = \frac{p(y|M_i)}{p(y|M_j)} \tag{5.30}$$

If the researcher decides to mix over all the possible available model, then *Bayesian model averaging* is used. Koop (2003) gives the point estimates of these average as:

$$E[g(y^*)|y] = E[g(y^*)|y, \, M_i]p(M_1|y) + E[g(y^*)|y, \, M_2]p(M_2|y) \tag{5.31}$$

where $g(.)$ is any function of interest. The posterior model probabilities are used as weighting factors.

Heretofore, we have focused less on the marginal likelihood but it is obvious that for a successful model selection exercise, the marginal likelihood has to be calculated. As explained earlier, the marginal likelihood is the normalising constant of the posterior density, which enables us to introduce the prior and the likelihood for each model. The calculation of the Bayes factor, posterior odds ratio, and the model averages all depend this quantity.

There are different methods of calculating the marginal likelihood, the two most common methods in Bayesian Econometrics are:

1. The Gelfand and Dey Method
2. The Chibs method

In this research, we use Chib's method of marginal likelihood calculation, this is because apart from the fact that the approach allows for calculation of the marginal likelihood in high dimensional parameter space, it is also easier. We refer the reader to Gelfand and Dey (1994), and, Koop (2003 pp. 105 - 106) for further explanation on the Gelfand and Dey method. We note that both methods are in the group of *the harmonic mean estimator*. Didelot et al. (2011) state that Chib method cannot be used when the likelihood function does not exist or is very difficult to evaluate. This, they say, is especially the case in the study of disease epidemiology and genetics.

Chib re-writes equation (5.24) as:

$$m(y) = \frac{p(y|\theta)p(\theta)}{p(\theta|y)} \tag{5.32}$$

Equation (5.32) is known as the *Basic Marginal Likelihood Identity* (see chapter 8 for a full explanation of chib's method).

We believe that the brief introduction to model selection as explained above will make the reader appreciate the Bayesian technique better and how 'easy' it is to carry-out. We will expatiate on these foundations in chapter 8. According to Murillo-Zamorano (2004), the use of Bayesian techniques in efficiency measurement enriches the researcher with more flexible models. He further states that the Bayesian technique also helps minimise the problem of imposing *a priori* distributional assumptions on the technical efficiency variable. That is, if the researcher has a vague knowledge of the distribution of the efficiency variable, it does not preclude estimation as the non-informative priors becomes handy here (though, the calculation of the marginal likelihood becomes difficult to execute).

5.4.2. Basic Differences Between Bayesian and Frequentist Econometrics

In table 5.2 below, we present the difference between the Bayesian and Frequentist methods with the sole aim of further enlightening the reader. Griffiths (1988), Poirier (1995), Bolstad (2007), Hajek (2012), are a good read for readers who are interested in further explanation in this area. Aside the reasons stated below, we repeat that the most important reason for the use of the Bayesian technique in this study is mainly due to personal preference and choice.

Table 5.2.: Basic Differences Between the Bayesian and Frequentist Econometrics

S/N	Bayesian Method	Frequentist Method
1.	They have a subjective view of probability	They have an objective view of probability
2.	Probability is interpreted as relative weights placed on parameters by each researcher called *prior probability*	Probability is interpreted has long-run relative frequency
3.	Statistics are based on the posterior distribution	Statistics are based on the sampling distribution
4.	The posterior distribution given the data renders complete inference about the parameter.	The sampling distribution gives different inferences about the parameter. This includes point and interval estimations and hypothesis testing
5.	The population parameter is assumed to be random and thus it is known	The population parameters is assumed to be fixed but unknown
6.	Probabilities calculated are post-data because the posterior distribution is found after the observed data has been taken into analysis	Probabilities calculated are pre-data because they are based on all possible random samples and not the specific random sample obtained
7.	Inferences are based on the actual data collected by the researcher	Inferences are based on all possible data that are likely to have occurred but did not.
8.	Inferences are based on Bayesian credible interval	Inferences are based on confidence interval
9.	It is easy to make predictive inference using this method	It is not easy to make predictive inference using this method
10.	It has an easy way to deal with *nuisance* parameters	It is not easy to deal with *nuisance* parameters
11.	Accuracy of the results are based on how much prior knowledge the researcher has about the parameter	The results are not affected by the prior knowledge of the researcher
12.	Bolstad (2007) states that most of the time they have smaller mean squared errors and thus they are closer to the population true value	They have larger mean squared errors most of the time and are farther from the population true value

5.4.3. Which is the First Paper Published on the Bayesian Analysis of the Composed Error model?

The literature as always (Kalirajan and Shand 1999 p. 166 and Murillo-Zamorano 2004, p.58 for example) referred to van den Broeck et al. (1994), and, Koop, Osiewalski and Steel (1994) as the first application of the Bayesian method in the analysis of the composed-error model. We investigate this assertion and after several searches on the internet and using the search terms "Bayesian, 'stochastic frontier"' and "Bayesian, 'composed-error"', it returned the work of Huang (1984) as the earliest attempt at applying the Bayesian technique to the composed-error model. Huang (1984) estimates the stochastic frontier model via the two methods -sampling theory and Bayesian theory - using the expected-maximization (EM) algorithm.

The production function used is of the form:

$$Z_i = \beta' X_i + v_i \tag{5.33}$$

$$Q_i = Z_i - U_i \tag{5.34}$$

where $i = 1, 2, \ldots, N$, Z_i is a random latent frontier output with half-normal distribution and v_i is a random error, X_i is a non-stochastic $k \times 1$ vector of inputs, Q_i is the observed output, which is related to equation (5.36) by equation (5.37), U_i represents technical efficiency, while, v_i is the double-sided random error. The model parameter vector is $\theta' = [\beta', \sigma_v^2, \sigma_u^2]$. The author makes use of a *diffuse* prior in the form of $1/(\sigma_v^2 \sigma_u^2)$ and the posterior density for θ is given as:

$$F(\theta|Y, Z, X) \propto (\sigma_v^2)^{-(n+2)/2} \exp\left[-1/2\sigma_v^{-2} \sum (Z_i - \beta' X_i)^2\right] \cdot (\sigma_u^2)^{-(n+2)/2} \exp[-1/2\sigma_u^2 \sum (Q_i - Z_i)^2] \tag{5.35}$$

In arriving at the estimates, the author applies the two-step procedure of Jondrow et al. (1982) to the model above. The posterior mode, $\hat{\theta}$, is obtained from the expected-maximization algorithm and the values of the stochastic frontier, Z_i, and the technical inefficiency, $\hat{U}_i$, are obtained by:

$$\hat{Z}_i = E(Z_i|Y_i, X_i; \hat{\theta}) \tag{5.36}$$

$$\hat{U}_i = \hat{Z}_i - \hat{Q}_i \tag{5.37}$$

He compares the results of the maximum likelihood and the Bayesian techniques and observes that they give similar estimates.

5.4.4. Gibbs Sampling and the Bayesian Approach to Stochastic Frontier Models

The advent of very fast computers in the early 90's and the development of the Gibbs sampling (Geman and Geman 1984; Tanner and Wong 1987; Gelfand and Smith 1990; Casella and George 1992) inspire a lot of work in the application of the Bayesian technique to stochastic frontier models. The Gibbs sampler is a technique for generating random variables from (conditional) distributions without calculating it from the density. Koop, Osiewalski and Steel (1994) describes the utilisation of the Gibbs sampler in stochastic frontier models. They utilized the same example used in van den Broeck et al. (1994) and found that Gibbs sampling resolves a lot of the computational issues encountered in their work.

The stochastic frontier model used is stated thus:

$$Q_i = X_i'\beta + v_i - z_i \tag{5.38}$$

They consider four different distributional specifications for the inefficiency term, z_i (for notational convenience, we replace u_i by z_i); three different Gamma distributions with shape parameters, j =1,2,3, and, one truncated normal distribution. The Gamma distribution is written thus:

$$p_j(z_i|\theta_j) = h_G(z_i|j, \lambda^{-1}) = \frac{z_i^{j-1}}{\lambda^j \Gamma(j)} \exp\left(-\lambda^{-1} z_j\right) I(z_j \geq 0) \tag{5.39}$$

The parameters of interest for the Gamma distribution are $\theta_j = (\beta', \sigma^{-2}, \lambda^{-1}, z_i)$; $I(.)$ is the indicator function; $\Gamma(.)$ is the Gamma function; and $h_G(.|j, \lambda^{-1})$ indicates the Gamma density with parameters j (j=1,2,3) and λ^{-1} and the truncated-normal distribution states thus:

$$p(z_i|\theta) = p(z_i|\delta\omega, \omega^{-2}) = \frac{h_N^1\left(z_i|\delta\omega, \omega^2\right)}{\Phi\left(\psi\right)} I\left(z_i \geq 0\right) \tag{5.40}$$

79

where the parameters of interest are $\theta = \left(\beta', \sigma^{-2}, \psi, \omega^{-2}, z_i\right)'$; $\Phi(.)$ is the cumulative distribution and h_N^1 is the normal density with mean defined as $\delta.\omega$ and variance ω^2.

For Gibbs sampling to work, the conditional distribution of the choice parameters must be in a closed form, in other words, they must be tractable; also, it must be possible to draw from this distribution. The authors arrive at the following conditional distribution for each of the choice parameters above; the conditional distribution of β is of the Normal distribution, for σ^{-2} is of the Gamma distribution, ψ is of no known distribution while the distribution of ω is not explicitly stated. The distribution for z_i depends on the value of j. If $j = 1$, equation (5.40) gives the exponential distribution, then z is of the univariate normal form truncated to be non-negative, the form of z_i is not known (that is, it is not in the closed form) when $j = 2$ or 3, hence, they used the importance sampling method to draw for samples, (though, in a later publication of the same work in Koop, Steel and Osiewalski 1994, they recommended the Metropolis-Hastings algorithm), to draw for z_i. Finally, for the truncated-normal case, z_i is of the truncated normal form. They showed the algorithm to follow in order to draw from the aforementioned distribution with (1-1)th pass. The Metropolis-Hastings algorithm is a generalisation of the Gibbs sampling. Its origin dates back to the work of Metropolis et al. (1953) and further improved by Hastings (1970), hence, the name Metropolis-Hastings.

Before Gibbs sampling became commonplace in stochastic frontier analysis, van den Broeck et al. (1994) utilize importance sampling as the method of drawing a sample from a distribution. They explain the concept of Bayesian model averaging, by this, they criticise the two step procedure of Jondrow et al. (1982) and Greene (1990). Koop, Osiewalski and Steel (1994) present the cost frontier with an asymptotically ideal model version of the composed error model with non-constant returns to scale. They replicate all of the distributional forms used in van den Broeck et al. (1994) and also addressed the issue of uncertainty relaxing the linear assumption used in van den Broeck et al. (1994). They utilize a semiparametric method in analysing their results and observe that estimates of the inefficiency term are sensitive to the choice of the functional form.

The Bayesian technique of measuring economic efficiency has also been applied to panel data with Koop et al. (1997) and Fernandez et al. (1997) serving as pioneering works in this area. Koop et al. (1997) utilize the stochastic cost frontier model for panel data and develop Bayesian analogue to the classical fixed and random-effects models under different classes of priors. They observe that the Bayesian fixed effects are characterised by the lack of links between the individual effects and the use of hierarchical prior is precluded. On the other hand, the Bayesian random effects model has a link between the individual effects and the hierarchical prior, hence, the use of a hierarchical prior is important. But McCulloch and Rossi (1994) remark that in the Bayesian point of view, there is no distinction between the fixed and random effects, only hierarchical and non-hierarchical models. This is because the parameters themselves are random (Holloway

et al. 2005). The individual effects are assumed to be constant over time. In other words, the economic efficiency measured is a time-invariant type of model. They apply their model to a panel of hospitals in the United States from 1987 - 1991.

Fernandez et al. (1997) investigate the use of improper priors in cross-sectional models. They emphasize that the use of improper prior in cross-sectional data does not necessarily lead to proper posterior distribution. Thus, they criticise use of improper prior in van den Broeck et al. (1994) and Koop et al. (1997). Fernandez et al. (1997) state that the use of certain improper non-informative priors in the composed - error analysis of a cross-sectional data will lead to improper, under-identified, and biased posterior estimates. One of the main reasons they put forward is that the number of parameters in the equation is greater than the total number of samples. In other words, we have $n + k + 2$ parameters to estimate; these are the n inefficiency variables$(z_1, z_2, \ldots, z_n)$, the k mean inefficiency parameter, λ, and two other parameters - β and σ - ; however, we have n observation points to carry out our analysis . They opined that this is however not the case if the data was longitudinal (panel) in nature (see Fernandez et al. 1997 for further explanation). They state that the use of panel data circumvents this problem because the researcher can impose a structure on the inefficiency terms (this last line seems to be at variance with Schmidt and Sickles (1984) where they state that imposing a structure on the inefficiency terms is not essential)

On the contrary, certain schools of thought counter the argument about the use of improper priors for cross-sectional data; they state that the advancement of the hierarchical modelling technique first introduced by Lindley and Smith (1972) circumvent this problem, as a result, the use of any type of improper priors does not lead to unbiased estimates. Their argument is further buttressed by Fernandez et al. (1997, p.14) where they state that "... the posterior distribution does not exist in the theoretical mode .... Of course, in the computer implementation of the empirical analyses, there will implicitly be truncation due to computer limitations. ... one could argue that the [improper] prior ... effectively becomes proper, thus, leading to a well-defined posterior distribution".

These said, it is important to carry out an empirical comparison of the classical and Bayesian techniques to efficiency measurement.

Kim and Schmidt (2000) compares the classical procedures such as multiple comparisons with the best, based on the fixed effects estimates; a univariate version "marginal comparisons with the best" bootstrapping of the fixed effects estimates; and maximum likelihood given a distributional assumption with the Bayesian procedures such as a Bayesian version of the fixed effects model, and various Bayesian models with informative priors for efficiencies. They apply these models to a large number of previously analysed datasets in order to obtain point estimates for technical inefficiency levels of firms and make inferences on these inefficiency levels by constructing confidence intervals on the point estimates. They observe that considerable differences do not exist between the classical

and Bayesian methods.

The Bayesian method of efficiency analysis has developed tremendously over the past decade and a half, such that, they have produced Bayesian equivalents to classical models of efficiency measurement.

Beyond this, they have gone ahead to resolve some intractable issues in the classical methods of efficiency measurement. Koop (2002) and Koop et al. (1999) use the Bayesian method to decompose output changes to technical, efficiency and input efficiency. Koop (2001), Fernändez et al. (2000) derive Bayesian tools for handling multi-output production frontiers. Fernández et al.(2002 and 2005) extends Fernändez et al. (2000) work by developing different definitions of efficiency in a multiple-output production system. This time they assume non-separability in inputs and outputs. They also introduce a situation where some of the outputs are undesirable. They test their propositions using two practical data – banking and agricultural data. They add a caveat saying care should be taken when deciding on which efficiency definition to use. They advise that researchers should look into theory for help on which definition to use. In the absence of such guidance the researcher should present results for several choices.

Tsionas (2002) presents a scenario where different firms in an industry face different technological frontier. Thus, they differ not in the inefficiency alone, but also on the frontier facing them. He, therefore, proposes a stochastic model with random coefficients to resolve this problem. He uses the Bayesian techniques to resolve this problem and draws inferences from it. He applies this method to a set of electric utility data. Also, Holloway and Paris (2002) utilize the Bayesian techniques to analyse the contentious von Liebig production frontier model. They develop an algorithm for it using the Markov Chain Monte Carlo (MCMC) algorithm. They resolve the issue of having an intercept higher than the yield plateau by imposing restrictions on the Gibbs algorithm.

O'Donnell and Coelli (2005) show how to impose the regularity conditions of monotonicity, quasi-convexity and convexity using panel data. They use the Bayesian technique to impose these conditions. Their work serves as a breakthrough in the efficiency literature because these conditions are difficult to impose. At present this condition is only imposed using the Bayesian methods. They adopt the Markov Chain Monte Carlo techniques of Gibbs sampling with data augmentation, and, the Metropolis within the Gibbs algorithm. They caution that their approach works when a time-invariant effect is assumed. They assume a truncated distribution for the inefficiency variable. However, they did not apply this approach to any agricultural data rather they use data from seventeen European railways. Their findings show that there are significant changes in the elasticities of the variables and shadow price ratios when the regularity conditions are imposed.

O'Donnell (2012) uses the Bayesian systems method to circumvent the problem of correlation between the explanatory variable and the error term in distance functions (He cites

Fernández et al 2000). He applies this method to state-level data on United States' agricultural input and output quantities. He calculates the distance function and the Total Factor Productivity (TFP) change. O'Donnell (2012) decomposes the total factor productivity into technical and a measure of efficiency change. He concludes that the primary driver of long-term productivity change in the United States is technical progress.

Also, O'Donnell and Griffiths (2006) analyse risk using state–contingent frontier. Since state of nature like risk is unobserved, they treat it like a latent variable. They then apply a Bayesian finite mixture model to analyse the state-contingent frontier of the Philippine rice data. The state-contingent framework gives different elasticities and technical efficiency estimates from the conventional frontier. The technical efficiency value is higher than the conventional frontier. This, they say, is because the error components in the conventional model measure noise, inefficiency and risk. In other words, the conventional frontier over-values noise and inefficiency.

Another important issue that the Bayesian method surmounts is the fact that all of the Bayesian papers reviewed above are able to calculate the finite sample properties of the firm-specific efficiency, including other variables of interest using 'easy' and tractable means which is not possible under classical method.

Sections (5.3) and (5.4) succeed in presenting the Frequentist and Bayesian methods of efficiency measurement. In the next section, we present the diverse applications of these methods in the literature.

5.5. Diverse Applications in the Literature

The stochastic frontier model has been utilized in diverse areas and we present some of these areas under different headings as follows:

5.5.1. Agriculture and The Environment

The parametric frontier analysis has received several applications in the area of agriculture in developing countries. Coelli et al. (2003) make use of the stochastic frontier to calculate the total factor productivity for a panel of crops in Bangladesh. The data contains thirty-one observations collected between 1960/61 and 1991/92 from 16 regions and the result reveals technical change is convex in nature with increase starting about the time of the introduction of the green revolution varieties in the 1970s. Technical efficiency reduces at an annual rate of 0.47 percent during the period they investigate. This has an effect on the total factor productivity which declines at the rate of 0.23 percent per annum with the rate of reduction increasing in later years. This, they say, raises questions about food security and the need to increase agricultural productivity in Bangladesh. They point out the non-use of price data in their analysis, which makes their work different from other authors (Coelli et al. 2003 cites Pray and Ahmed, 1991, and, Dey and Evenson, 1991).

Wadud and White (2000) compare the stochastic frontier approach with the data envelopment analysis and discover both methods indicate efficiency is significantly affected

by irrigation and environmental degradation. Amaza and Olayemi (2002) investigate the technical efficiency of food crop farmers in Gombe State, Nigeria and arrive at similar mean technical efficiency as Ajibefun and Abdulkadri (1999). However, the difference between the minimum and maximum technical efficiency score for Amaza and Olayemi (2002) is seventy-six percent, while for Ajibefun and Abdulkadri (1999) is about sixty-six percent.

Jara-rojas et al. (2012) examine the impact of the adoption of soil and water conservation practices on productivity. They find a positive relationship between soil and water conservation and technical efficiency. They find that an enhancement of the technical efficiency also improves the net returns on investment. Reinhard et al. (1999) estimate the technical and environmental efficiency of a panel of dairy farms. They assume the production of two outputs – dairy and excessive use of Nitrogen. They analyse their efficiencies separately. Their objective involves investigating whether farmers can both be technically and environmentally efficient. They also examine the compatibility of these two types of efficiencies. They estimate a stochastic translog production function to obtain output-oriented technical efficiency. On the contrary, they estimate input-oriented technical efficiency for the Nitrogen surplus of each farm. They obtain a mean output-technical efficiency of 0.894 while the input-oriented environmental efficiency is 0.441. They remark that intensive dairy farming is both technically and environmentally more efficient than extensive dairy farming.

Reinhard et al. (2000) also examine comprehensive environmental efficiency in Dutch dairy farms. This paper is a continuation of Reinhard et al. (1999) paper. In this paper, apart from surplus Nitrogen which they use in their earlier work, they also investigate excessive use of phosphate and total energy use of these farms. They compare efficiency scores in the stochastic frontier analysis with the data envelopment analysis. The mean technical efficiency values for the two methods of analysis are different. The stochastic frontier has an output technical efficiency value of eighty-nine percent, while the data envelopment analysis has an efficiency value of seven-eight percent. There is a significant difference between their environmental efficiencies also. The stochastic frontier analysis records a value of eighty percent while the data envelopment analysis records a value of fifty-two percent. It is evident from the result of the two efficiencies that the stochastic frontier method over-values efficiency scores.

Kurkalova and Carriquiry (2003) investigate a small sample of forty-one Ukrainian collective farms from 1989 to 1992. They posit that since the data size is small, the number of years precludes the estimation of the data by the classical approach. They further justify the use of the Bayesian approach by the fact that they are interested in changes in the efficiency values at the farm-level in the use of different inputs. Also, they point to the fact that the precision of the estimate of the Metropolis-Hastings may be heterogeneous across firms. This, they affirm, will be difficult to investigate using the maximum-likelihood ap-

proach. The median of the posterior distribution of the average technical efficiency gives a value of 0.942.

Using rice farmers in the Philippines Dawson et al. (1991) measure farm-specific technical efficiency over time of these farmers. The range of efficiency values across the twenty-four farms in the survey is between eighty-four percent and ninety-five percent. They affirm that it will be difficult to relate the narrow spread of farm-specific inefficiencies to farm socioeconomic factors. Instead, they suggest that increase in rice production in the future will come from technological progress. They specify a Cobb-Douglas production function.

Heshmati et al. (1995) investigate technical efficiency, technical progress, and, the bias in technical change using panel data from the Swedish pork industry . The panel covers the period from 1976 to 1988. They use a generalized Cobb-Douglas model where input elasticities are linear functions of time. In other words, they use a time varying model to estimate inefficiency. They use farm dummies to capture farm heterogeneity. Elasticities with respect to animal and material inputs increase over time. They record a mean technical efficiency value of about ninety-one percent. About three percent of farms are inefficient, which shows most of the farms in the industry are efficient. They find technical change to be positive, but, regresses slowly between the periods 1981 to 1988. They observe that maize production is done under increasing returns to scale. Many of the maize farmers are technically inefficient. The mean technical efficiency he records is fifty-three percent. Farm-specific efficiency is as low as three percent, while, the modal efficiency class being about sixty percent. He observes that plot size, hired labour, use of hybrid seeds, education affect farmers technical efficiency. However, adoption of new technology does not have an effect on efficiency.

Ogundele and Okoruwa (2006) divide rice farmers in Nigeria into two groups: those that plant improved variety and those that plant local variety. Afterwards, they compare their technical efficiencies. Their results show there is no significant difference in the technical efficiency of these two groups. They assert that farming experience and number of visits by extension agents are the only socioeconomic characteristics where significant differences between the groups exist. They observe that farm size was the most important factor that influences technical efficiency in Nigeria. They query the success of the various rice development programmes in Nigeria as rice farmers' technical efficiency is still low.

Chirwa (2003) examines the sources of technical efficiency among maize farmers in Malawi. His result show that the farms have an average efficiency score of about fifty-three percent, while, fifty – eight percent of the farms have efficiency scores below sixty percent. We can then infer that inefficiency is still high in Malawi among maize farmers. They observe that inefficiency falls with plot size, use of hired labour, use of hybrid seeds and membership of a farming association.

Also, Amos (2007) analyses the productivity and technical efficiency of smallholder cocoa farmers in Nigeria. He uses a Cobb–Douglas stochastic production function. He utilizes

primary data of two hundred and fifty cocoa farmers in Ondo State, Nigeria. He shows that the farmers are experiencing increasing returns to scale in resource use. The average farmer needs to increase his efficiency by about twenty-eight percent. However, the difference in efficiency values between the least efficient farmer and the most efficient farmer is eighty percent. This is extremely high. He explains that age, education, and, family size influence technical efficiency. He advises the government to introduce a sustainable education policy in Nigeria.

Binam et al. (2004) use a Cobb-Douglas production function to examine factors affecting technical efficiency among smallholder farmers in Cameroon. He surveys four hundred and fifty farmers over fifteen villages in the 2001/2002 planting season. They calculate the technical efficiency for each crop (groundnut and maize) and then for joint production of groundnut and maize. They obtain seventy-seven percent, seventy-three percent and seventy-five percent respectively as efficiency values for groundnut monocrop, maize monocrop and joint production of groundnut and maize. They conclude that the differences in efficiency are as a result of the credit, soil fertility, social capital; distance of the plot from access roads and extension services.

Alene and Hassan (2006a) study the efficiency of the intercropping systems of production in Ethiopia . They examine annual and perennial crops in southern Ethiopia. They analyse their data using the stochastic frontier analysis, data envelopment analysis and the parametric distance function. They compare the results of the stochastic frontier analysis, data envelopment analysis and the parametric distance function. They affirm that the efficiency rankings using the three different methods are similar. However, the single output stochastic frontier analysis records the lowest efficiency scores among the three methods. Results from the multiple output methods - the data envelopment analysis and the parametric distance function – are similar. The data envelopment analysis and the parametric distance function also show significant correlation with each other. Their efficiency estimates are bigger than those of the stochastic frontier analysis. The average technical efficiency is about ninety percent. This means that intercropping improves technical efficiency. They further remark that increase in productivity occurs from diversification from the subsistence method of food production to the cash crop method of food production. They comment that their findings are in line with Schultz (1964) work.

5.5.2. Public Policy and Evaluation

In the area of policy evaluation and analysis has also received a considerable amount of focus in the area of productivity and efficiency analysis. For example, Seyoum et al. (1998) use the Battese and Coelli (1995) stochastic production function to compare between farmers that participate in the Sasakawa-Global 2000 project and those who do not in Ethiopia. They collect twenty samples from two different districts (Keresa and Kombolcha) of eastern Ethiopia and show the difference in the levels of production in these two

districts by use of a dummy for one district. The data is panel in nature which justifies their use of the Battese and Coelli (1995) model. Seyoum et al. (1998) recommend that policy makers should expand the Sasakawa-Global 2000 project as farmers who participate have better output, higher productivity and greater efficiency than farmers who do not.

Still on the impact of government programmes on efficiency, Abdulai and Huffman (2000) look at the impact of the Structural Adjustment Programme on the efficiency of rice farmers in Northern Ghana using a stochastic profit function. Their results show rice producers in the region are highly responsive to market prices for rice and inputs. They support the introduction of the structural adjustment programme because it makes the farmers more market oriented. Also, Ajibefun and Abdulkadri (1999) find the Cobb-Douglas production function as being adequate to represent the efficiency of Nigeria's National Directorate of Employment Farmers (NDE) Scheme. They reject the half-normal distribution assumption for the inefficiency term. Ajibefun (2002) simulates the impact of policy variables on the technical efficiency of small-scale farmers in Nigeria. He discovers that increase in education level and farming experience would significantly improve small-scale farmers' technical efficiency.

Kalaitzandonakes and Dunn (1995) investigate the effect of education on technical efficiency in eighty-two corn family farms in Guatemala. These families participate in government's market-based land reform programmes. They analyse the data using the deterministic, stochastic, and, data envelopment analysis. They compare the results of the deterministic frontier with that of the stochastic frontier. They find that the efficiency values in the deterministic frontier were higher than that of the stochastic frontier. On the contrary, most of the family farms are seen to be more efficient under the stochastic frontier than under the deterministic frontier. Comparison of the results of the data envelopment analysis with the other two frontiers suggests technical inefficiency is minimal. They support the assertion that empirical technical inefficiency measurements are not perfect measures of their latent theoretical analogue.

Some researchers investigate the impact of a particular economic reform on efficiency. Abdulai and Eberlin (2001) look at the effect of the economic reform in Nicaragua on maize and beans farmers' technical efficiency. They collect their data from two regions in Nicaragua during the 1994-1995 planting season. They employ the translog stochastic frontier model in analysing their model. Their analysis gives an average efficiency value of about seventy percent and about seventy-four percent for maize and beans respectively. Schooling, formal credit and farming experience have positive effects on efficiency while farmers' participation in non-farm activities reduces efficiency.

Areal et al. (2012) include the milk quota system as one of the factors affecting efficiency. They affirm that the way a farmer behaves in the milk market has links with his technical efficiency. They note that farmers who purchase or lease their milk quota are more

efficient. In the same vein, farmers who tend to go above the allocated quota are more efficient. They comment that other environmental production has links with technical inefficiency.

5.5.3. Livelihood and Fisheries

Most of the notable studies in the area of livelihood and fisheries use the Bayesian technique for efficiency analysis. Bezemer et al. (2005) examine the role of livelihood strategies in rural growth and poverty reduction in Georgia. The study incorporates livelihood diversity into the stochastic frontier. The results show that animals contribute more to output. They remark that involvement in non-agricultural activities have a positive contribution to efficiency.

Flores-Lagunes et al. (2007) use the results of Horrace (2005) - he develops probability statements, ranks, and, selection rules for independent truncated-normal populations - to estimate the technical efficiency of thirty-nine fishing vessels in the Northeast Atlantic herring fleet on the basis of each vessel's probability of being efficient. They develop a selection technique to identify groups of vessels at a pre-stated probability levels.

Holloway et al. (2005) present a pedagogic way to estimate the composed error model hierarchically. They explain how the marginal likelihood is calculated in a composed error model. They show how this procedure works by use of data from the Pacific Coast Groundfish Fishery. They also introduce the "good captain" hypothesis in their analysis. " The good captain" hypothesis focuses on the fact that skill and talent of the fisherman and not luck bring large harvest to the fisherman.

Alvarez et al. (2003) working on hake catches in Northern Spain disagree with this hypothesis. They demonstrate that luck rather skill is more important. We note that Alvarez et al. (2003) use the classical method. Holloway and Tomberlin (2007) take the work of Holloway et al. (2005) a step further by calculating efficiency estimates as well as the probabilistic rankings of the relative technical efficiency of fishing boats. The sample is made up of ten thousand eight hundred and sixty-five fishing boat trips in the United States Pacific hake (or whiting) fishery during the period 1987 to 2003. They make use of the Gibbs sampler in arriving at their conclusion. They explore the likelihood of a particular boat being 'best' simultaneously with the technical efficiency scores. They ignore the problem of heterogeneity which results from different boats having different production frontiers.

Tomberlin and Holloway (2008) further present a Bayesian hierarchical approach to estimating the composed-error model in fisheries. They use the same sample as in Holloway and Tomberlin (2007). They estimate two cross-sectional models for comparison with panel models. They divide their data into two layers of hierarchical models in each case. They discover that doing this yields the best results in the cross-sectional and panel data models considered. The results are similar in both models, though their marginal like-

lihood differs a bit. They make two salient conclusions. The first is that hierarchical panel models that allow for boat-specific and year-specific efficiency measures are better than less elaborate specifications. Secondly, they observe that there may be a progressive outward shift in the efficient frontier in the study period.

5.5.4. Health

The use of the stochastic frontier model to estimate the effect of health on farmers' efficiency receives special focus in the literature. Croppenstedt and Muller (2000) take up this challenge when they research into the role of the Ethiopian farmers' health and nutritional status on their productivity and efficiency. They find that distance to the source of water as well as nutrition and morbidity affect agricultural productivity. Surprisingly, elasticities of labour productivity regarding their nutritional status are strong. They further affirm that this strong correlation continues with technology estimates and wage equations. However, they record considerable loss in production due to technical inefficiency even after accounting for health and nutrition of workers.

Ajani and Ugwu (2008) look at the impact of adverse health on the productivity of farmers living in the Kainji basin of North-Central Nigeria . Their study shows the health variable as being positive, large, and, statistically significant. They, therefore, conclude that health capital is an essential input in agriculture. Ajani and Ashagidigbi (2008) use numbers of days of incapacitation as a proxy for malaria incidence in Oyo State, Nigeria . Surprisingly, they ran a normal linear regression to investigate the effects of malaria on agricultural productivity. Their analysis shows that age and the days of incapacitation are insignificant statistically. Their result may be different if they use a stochastic frontier.

Olarinde et al. (2008) explore the factors that affect beekeepers' technical efficiency in Oyo state, Nigeria. They observe that the beekeepers are efficient by about eighty-five percent. In other words; there is still room for them to increase their efficiency by fifteen percent. They observe that one major determinant of efficiency in their study is the status of the bee-keepers, that is, if they are full-time or part-time bee-keepers. Marital status is also another variable that affects technical efficiency, they note. Thus, a farmer who is single is likely to be more efficient than a married farmer.

Apart from Croppenstedt and Muller (2000), another researcher who attempts to explore the impact of farmers' health on agricultural productivity is Loureiro (2009). She uses secondary data from Statistics Norway and Tax Revenue Service of Norway to investigate this. She uses the heteroscedastic stochastic frontier model. She includes the farmers' health as one of the variables that affect agricultural efficiency. She looks at health as anything that could cause physical injury and hazards to the farmer. Her result shows the farmers' health plays statistical significance in affecting his efficiency. She recommends that government should endeavour to enlighten the farmers on health and safety attitudes on the farm. In the same vein, Ulimwengu (2008) studies the effect of health on farmers' technical efficiency. He estimates a Cobb-Douglas stochastic frontier to explore this

relationship. He uses data from the fifth round of the 1999 Ethiopian Rural Household Survey (ERHS). He observes that illness increases the probability of child labour. He motivates his conceptual framework from the household model, modifies and carries out comparative-statics analysis on it. He further asserts that reducing the remoteness of villages will lessen the likelihood of the farmer being constrained by illness. He infers that sickness reduces agricultural efficiency.

5.5.5. Spatial Studies and Migration

There could be spatial differences in the technical efficiencies of different farms based on ecological differences, farm size and interactions between these two variables. Tadesse and Krishnamoorthy (1997) set out to investigate this in their research on paddy rice farmers in the state of Tamil Nadu, India. They remark that the farmers still have the opportunity of increasing their efficiency by seventeen percent. They observe significant variation in the mean technical efficiency in the four zones that make up Tamil Nadu. They use a two stage approach where the first stage is to obtain farm-specific technical efficiency and then use a Tobit model to compare the differences in the technical efficiencies of each region and zone in the second stage.

Wang and Schmidt (2002) note a bias in the results obtained by this process and they went ahead to use the Markov chain Monte Carlo technique to prove that there is a serious bias at every stage of the procedure. Chen et al. (2009) also examine the technical efficiency of farms in four regions of China. The four regions are North, North-East, East and South-West. They observe that different inputs need to be put to efficient use in the different regions. For example, inefficient use of industrial input is the main problem in the East while in the North it is capital. They assert that farms in the North and North-East are relatively more efficient than farms in the East and South-West. They recommend a change in the land tenure system to eliminate land fragmentation in China.

Mochebelele and Winter-Nelson (2000) examine the effect of migratory labour to mine fields on farm technical efficiency in South Africa. They try to establish if migrant labour actually complements farm production or not. They establish that households with migrant farmers have higher production and are more efficient than households without migrant farmers.

5.5.6. Mixtures of techniques and/or Models

We discuss this group of studies separately because there are a few research that attempt to either compare the parametric and non-parametric methods, or, examine the different types of efficiency.

A paper that successfully combined the non-parametric technique of data envelopment analysis and an econometric model is Audibert et al. (2003). They use a combination of the data envelopment analysis and the Tobit model to infer on the social and health determinants of the efficiency of cotton farmers in Northern Côte D'Ivoire. They use

the high density of the malaria parasite in the blood of an individual as a proxy for the health of the household. They use a two step process; firstly, they use the data envelopment analysis to arrive at relative technical efficiency values and then they regress this efficiency score against factors they think will affect the efficiency. The 'high density of malaria parasite in the blood' variable enters the model at the second stage. Their results show that malaria greatly reduces farmers' technical efficiency. They further conclude that it is intensity of infection by the disease that is more important rather than its presence. Our research collects data on the prevalence of the disease in an area rather than just hospital reported cases; this we believe gives further credence to our results.

Alene and Hassan (2006a) add to the few studies that attempt to examine the technical, allocative, and, economic efficiencies of farmers . They measure the efficiency of traditional and hybrid maize producers in eastern Ethiopia. Furthermore, their work inputs-returns-to-scale variable in the traditional efficiency decomposition method. They, then, compare their results with the conventional efficiency decomposition approach. They observe that the conventional efficiency decomposition approach over-states the true efficiency measures under increasing returns to scale production of hybrid maize. Also, it under-states efficiency figures under decreasing returns to scale production of traditional maize. They further observe that the mean of technical and economic efficiencies are considerably lower in the conventional method than the scale-adjusted efficiency measure in traditional maize production and higher in hybrid maize production. They affirm that if one uses the conventional method one will draw the wrong inference. The reason, they say, is that one is likely to say traditional maize production exhibits significant inefficiencies while hybrid maize cultivation exhibits lower inefficiency.

Rios and Shively (2005) use a two step approach to examine the technical and cost efficiency of coffee farmers in Vietnam. The first step involves the use of the data envelopment analysis to examine the technical and cost efficiency of the farmers. They estimate a Tobit model in the second step to explore the factors that affect these efficiencies. They observe lower technical and cost efficiency on small farms. They suggest that inefficiency may not necessarily be caused by farm size. They point to factors like irrigation pipe length, higher education, access to credit and land tenure system as the likely factors that affect inefficiency.

Balcombe et al. (2006) apply the Bayesian, Classical stochastic frontiers and the data envelopment analysis to a sample of Australian dairy farms . They adopt van den Broeck et al.(1994)'s informative prior of 0.875 as their (in)efficiency value. They investigate the impact of imposing regularity conditions on their results at three different points in the data. Firstly, without regularity conditions imposed, secondly, with regularity conditions imposed at sample means and thirdly with regularity conditions imposed at all data points. They adopt the use of the random-walk Metropolis-Hastings step in imposing the regularity conditions at all points in the data. They emphasize that the imposition of the

regularity conditions does not change the efficiency results significantly.

Some researchers attempt to investigate if differences exist between two parametric techniques. For example, Shah et al. (1994) examine crop specific technical efficiency in Pakistan. The crops they investigate are wheat, maize, sugar cane and vegetable. They use the corrected ordinary least square approach and maximum - likelihood techniques to analyse their data. Both techniques show the farmers as being inefficient. Maize and Sugarcane give the highest inefficiency values among the crops. There are differences in the cause of inefficiencies among the crops. Their analysis show inefficiency in maize and sugarcane is due to technical inefficiency. On the other hand, in wheat and vegetable inefficiency is due to random shocks.

In the next section, we discuss the summary of this chapter.

5.6. Summary

In this chapter, we perused the literature and have attempted to elucidate thematic issues in the literature concerning efficiency analysis. Though, the Bayesian method has proved useful in resolving some issues on the stochastic frontier analysis (including the treatment of risks and uncertainty), but there is still a lot of issues to be debated on the application of the efficiency concept. One of the issues being debated is; can a variable affect both production and efficiency at the same time? From the aforementioned, it is obvious that the literature will continue to observe new developments on the composed error model both on the theoretical and empirical ground.

So far, we have presented the primal of the stochastic frontier production model, for a thorough explanation on the dual cost frontier the reader should see Christensen and Greene (1976), Schmidt and Lovell (1979, 1980), Greene (1980), Kumbhakar (1997), Greene (2008, p. 187) and Chambers (1988). For a thorough explanation on the problems associated with estimating this class of model the reader should see Melfi (1984), Bauer (1985 and 1990), Atkinson and Cornwell (1993 and 1994).

For comprehensive and thorough understanding of the estimation and decomposition of the profit frontier using the dual panel data frontier techniques, we refer the reader to Cornwell and Schmidt (2008) and Kumbhakar and Lovell (2000, pp. 60-61).

Another method of measuring efficiency that receives attention in the literature is the Stochastic Varying Coefficients Frontier Approach. This method uses the varying coefficient approach of Swamy (1970). For a full introduction into this technique see Kalirajan and Obwona (1994), and, Kalirajan and Shand(1999, pp. 164).

Apart from the work of Strauss (1986), who investigates the effect of nutrition on labour productivity, and, Fernändez et al. (2005), who assumes non-separability in input and output, the literature is scarce with research that effectively integrates these two models in their analysis. Our research attempts to reduce the void in these areas of research.

In the next chapter, we review the literature on the Markov chain Monte Carlo diagnostics, while, in chapter 7, we peruse through the data, how it was collected, and, how these affect our estimation procedure. After this, we can then seek to value equation (4.13) in chapter four in empirical terms and see where the application of the Bayesian methodology comes in.

6. Markov chain Monte Carlo Convergence Diagnostics (A Review)

6.1. Introduction

In the last chapter, we reviewed the literature on the stochastic frontier model. In this chapter, we attempt to peruse the literature once again on Convergence Monitoring. Also, in the concluding part of this chapter, we state that we use the autocorrelation and the Raftery and Lewis methods to test for convergence in our models. The selection of this two methods was by choice of the researchers.

The chapter begins with theories that form the foundation of Markov chain Monte Carlo convergence monitoring and proceed to emphasize that the true definition of the term is a grey area that researchers need to reach a consensus upon. We then go ahead to espouse the different convergence monitoring procedures available to Bayesian econometricians. However, before we go on with our exposition we define two key terminologies that the reader will encounter in the literature. These are Markov-chain Monte Carlo and burn-ins. Markov chain Monte Carlo are a type of simulation in which a draw (replication) is dependent on previous values; this is the essence of conditioning. Examples of this type of simulation are Gibbs sampling and Metopolis-Hastings sampling (we espouse these two methods in chapter eight). Burn-ins are values which are set to enable the distribution stabilize at the initial stage of drawing for samples. At this stage, the draws are irregular and tend to draw from region of low density. The burn-in gives it time to stabilize in order to draw from high density regions.

6.2. The Foundation Theories

The term convergence draws from two important theories in the literature, which are, the Weak Convergence theory and the Central Limits theory. There are several theorems in the literature based on these two theories, a good read is Billingsley (1971), Garśia (1970), and, Van Der Vaart and Wellner (1996).

Billingsley (1971) states the weak convergence theory thus: suppose that S is a metric space of open sets where a *Borel-set*, Υ is generated (the term Borel-set is a standard term in statistics we refer the reader to any analytical textbook in statistics for further explanation). Assuming P is a completely additive set function with nonnegative values

on Υ such that $P(S) = 1$, and P and P_n (where n are nonnegative values) are probability distributions on (S, Υ), it is said that P_n converges weakly to P (written $P_n \Rightarrow P$), if

$$\lim_{n \to \infty} \int f \, dP_n = \int f \, dP \tag{6.1}$$

for all functions f in a class of bounded, continuous real-valued function on S denoted by $C(S)$.

He states that the Central Limit Theory holds when:

$$\frac{1}{\sigma \sqrt{n}} S_n \Rightarrow N \tag{6.2}$$

where σ is a positive constant and N is a normally distributed random variable with mean 0 and variance 1; n and S are as defined earlier.

We are quick to note that these two theories do not tell a researcher when convergence to the stationary distribution will occur. This necessitates the development of methods by which convergence can be monitored in the literature, however, there are thematic issues in the literature which we espouse in section (6.3) below.

6.3. Thematic Issues

The Gibbs sampling and the Metropolis-Hastings methods are a class of Monte Carlo sampling techniques called Markov chain Monte Carlo methods. One major problem with this class of sampling methods is deciding when to stop sampling, that is, what amount of draws will just be enough before one can say the samples are representative of the underlying stationary distribution. This is the general view of the word *convergence.* The exact definition of the word itself and its meaning depends on the researcher and what diagnostics he decides to use. This is because as stated in Cowles and Carlin (1996), this definition of the word *convergence* is just for convenience and to serve as a point to begin the discussion on this important subject. Brooks and Roberts (1998,p.5) seem to corroborate this point when they state that the choice of convergence diagnostics by the researcher is problem specific and that some convergence diagnostics can only be used for certain types of algorithm.

They state that the nature of the Markov chain itself shows that what is produced by this process itself is a "sample from a distribution". As a result, it does not exactly fit into this definition, as each draw from a Markov chain is dependent on the previous draw and this causes an unnecessary delay in the ability of the algorithm to draw from the stationary distribution (see Cowles and Carlin 1996 p. 883).

We have highlighted the thematic issues raised in the literature. We proceed to expatiate on the two different approaches - the theoretical and diagnostic approaches - of monitoring convergence.

6.4. Approaches to Monitoring Convergence

Cowles and Carlin (1996) state two approaches to monitoring convergence. These are the (i) Theoretical and (ii) Diagnostic Approaches to monitoring convergence.

The theoretical approach of measuring convergence includes the methods of Polson (1993), Rosenthal (1993, 1995a and b), Didelot et al. (2011,p.71-72). One major disadvantage of this method is the amount of mathematical calculations that needs to be done by the researcher for every model used. This disadvantage endears the applied econometricians to the diagnostic approach. Ripley (1987) refers to them as them as *output analysis* techniques.

Lesaffre et al. (2012) identify different types of convergence diagnostics which are divided into two main groups:

- The Graphical Approach
- The Formal Diagnostic Tests

In the next section we explain the different procedures associated with each of these two groups.

6.4.1. The Graphical Approach

The graphical approaches involve the use of plots to assess if the distribution has converged. These include the trace plot, autocorrelation plot, running mean plot, Q-Q plot, CuSum plot, and cross-correlation plot.

The Trace Plot

This includes the *thick felt-tip pen tests* of Gelfand and Smith (1990). It involves a simple visual assessment of the plot in order to assess the characteristics of the Markov chain. It involves the researcher taking a rigorous look at the plot from the beginning to the end of the draw in order to assess if convergence as taken place. The trace plot may be produced for each parameter in the distribution in order to assess the chain. It could also be a joint assessment of the whole chain. The researcher normally monitors the log of the likelihood function or the posterior distribution. If convergence to the stationary (target) distribution occurs then, the trace appears almost as a horizontal strip. In the case, of the thick felt-tip pen tests, convergence to the stationary distribution occurs when the trace

can be covered by less than the width of a felt-tip pen. Lesaffre and Lawson (2012, p.175) states that the trace plot can also show how quick the chain draws from the posterior distribution, they refer to this as the *mixing rate* of the chain They state that the "thick felt-tip pen" test is still popular in the literature. Cowles and Carlin (1996) states that this method does not work well in slow mixing chains, also, it leads to wastage of a lot of pre-convergence samples. Next, we discuss the autocorrelation plot.

The Autocorrelation Plot

Autocorrelation does not measure convergence at all, although, it is an indication of the mixing rate in the distribution and the statistical efficiency of the estimates. The correlation between lags in the distribution is usually assessed using the *Pearson correlation coefficient* or any time series approach. What is shown on the autocorrelation plot is the autocorrelation function. As stated in Lesaffre and Lawson (2012), when the autocorrelation decreases slowly with increasing lag, the mixing rate is low. Once convergence to the stationary distribution is attained the autocorrelation function does not change anymore, no matter the size and value of the autocorrelation. They state also that, when all autocorrelation is close to zero, then the Markov chain Monte Carlo sampling is done in an almost independent fashion and stationarity is close to zero. Unfortunately, the literature states that autocorrelation plot is not a reliable tool for assessing convergence. This is because a relatively high autocorrelation function does not necessarily mean convergence has not been obtained.

The Running Mean Plot

The running mean plot is also known as the *ergodic mean* plot. The running mean is the mean of all sampled values, including the starting value denoted by say t. Once convergence is attained the mean of the distribution shows stationarity beyond the point of convergence. Lesaffre and Lawson (2012) state that the initial variation in the running-mean plot values is high, but it steadies as the chain continues.

Q-Q Plot

This involves dividing the stationary distribution into two and plotting one-half on the x-axis and the other on the y-axis, such that any deviation from the central line is seen as nonstationarity in the chain. This plot emphasizes the point that a stationary distribution is stable notwithstanding at what point one decides to break the distribution.

The CuSum Plot Method

The Cumulative Sum (CuSum) methods monitor convergence using the CuSum plot of the sampler output. This method was proposed by Yu and Mykland (1998) with further exposition by Yu et al. (1995). The CuSum is presented formally thus:

Let X_0, X_1, ..., X_n be the output of a Markov chain with n_0 burn-in time and let θ be the summary statistics, then the observed CuSum or partial sum is given as:

$$\hat{S}_t = \sum_{j=n_0+1}^{t} [T(X_j - \hat{\mu})] \tag{6.3}$$

for $t = n_0 + 1, ..., n$

where $\quad \hat{\mu} = \frac{1}{n-n_0} \sum_{j=n0+1}^{n} T(X_j)$

Plot $\{\hat{S}_t\}$ against t for $t = n_0 + 1, ..., n$ and the successive points are connected by line segments (*explanation continues in the appendix*).

Yu and Mykland (1998) presented a subjective view of the CuSum plot while Brooks (1998) gives an objective view of the CuSum plot.

Cross-correlation plot

This is a scatterplot of two different sampler output. Lesaffre and Lawson (2012) state that this plot helps to indicate if the model parameters are strongly related and it is a good diagnosis for overspecified models.

The *Are We There Yet (AWTY)* method

In 2008, Nylander et al. develop a method of graphical exploration method for detecting convergence in *Phylogenetics* (the study of the history and development of organisms) by using the *Are We There Yet (AWTY)* program. The program takes as input a number of phylogenetic trees generated as output by other phylogenetic MCMC programs. Then, a number of diagnostic tests are now carried out on this output and viewed graphically. However, we would like to say that this method is not a new method of convergence diagnosis *per se* as it depends on other diagnostics to function properly. Also, the authors state that their method cannot guarantee the convergence of a chain; it is only meant to state when convergence has not occurred in a chain.

In practise, researchers often combine the graphical plots with other formal diagnostic tests. As a result, we discuss the formal diagnostic tests in the next section.

6.4.2. The Formal Diagnostic Approach

These tests are broken down using a modification to the headings used by Brooks and Roberts (1998). Apart from their work, the work of Cowles and Carlin (1996) have formed the basis of this section.

Variance Ratio Methods

This consists of the Gelman and Rubin (1992), the Brooks and Gelman (1998), and theBrooks and Giudici (2000) methods. Cowles and Carlin (1996) states that the procedure is based on normal approximations to exact posterior inference. They state that it consists of two steps, step 1, is to be carried out before sampling begins and it involves obtaining an overdispersed estimate of $m \geq 2$ sequences of length $2n$, with different starting point. In step 2, the first n iterations are discarded and the last n iterations are retained

and these are used to re-estimate the stationary (target) distribution of the scalar quantity using a conservative Student t distribution (*see appendix for further explanation*). A strong criticism of the Potential Scale Reduction Factor is that it heavily relies on the researchers' ability to find a starting distribution that is overdispersed with respect to the target (stationary) distribution. Also, relying on the normal approximation to the posterior distribution in diagnosing convergence to the true distribution may be questionable (See Cowles and Carlin 1996 for further elucidation on these).

Brooks and Giudici Methods:

Brooks and Giudici (2000) develop a method of monitoring convergence through the use of the Analysis-of-Variance. The method is a model selection method. The method is further development of the Gelman and Rubin (1992), and, the Brooks and Gelman (1998) methods. This involves the splitting the variation present in a Markov chain output into two independent groups. They explained that their method is basically focused on assessing convergence of complex model choice problems based on the reversible jump Markov chain Monte Carlo algorithm and the finite mixture models. Their method attempts to assess the performance of these models during simulation and assess the internal performance within the chain. The model also attempts to assess the influence of the starting point on the Markov chain Monte Carlo output.

We believe that this model is more tilted towards determining what causes complex models not to converge rather than monitoring convergence of the whole distribution. Also, Brooks et al. (2003) believe that the choice of parameters to use for this method might be difficult, also, monitoring the parameters alone might not reveal everything going on within the chain.

Spectral Methods

This diagnostic test method makes use of the Spectral Analysis to gain variance estimates via the spectral density $S(\omega)$.

Geweke Spectral Density Diagnostic:

Geweke et al. (1991) suggests testing for stationarity by dividing the Markov chain for θ^t into an early and a late part and comparing the means of the two distributions. If we denote the early part of the distribution as $\{\theta^t : t = 1, 2, ..., n_A\}$ and $\{\theta^t : t = n^*, ..., n\}$ where $1 < n_A < n^* < n$ and $n_B = n - n^* + 1$. Then their respective (posterior) means is given by $\bar{\theta}_A$ and $\bar{\theta}_B$. If we let $\hat{S}_\theta^A$ and $\hat{S}_\theta^B$ denote consistent spectral sensitivity estimates for n_A and n_B (Ripley 1987 explains further). If the ratios n_A/n and n_B/n are fixed width $\frac{n_A + n_B}{n} < 1$

Then, by the central limit theorem, the distribution of the diagnostic approaches a standard normal distribution as n tends to infinity. This is given by:

$$Z_n = \frac{(\bar{\theta}_A - \bar{\theta}_B)}{\sqrt{\frac{1}{n_A}\hat{S}_\theta^A(0) + \frac{1}{n_B}\hat{S}_\theta^B(0)}} \to N(0,1) \quad as \ n \to \infty \tag{6.4}$$

We may use the result in equation (6.4) to test the null hypothesis of equal location of $\bar{\theta}_A$ and $\bar{\theta}_B$, if it is not rejected it means that the distribution has converged. Geweke et al. (1991) suggested setting $n_A = .1n$ and $n_B = .5n$. Cowles and Carlin (1996) state that he attempted to resolve the issue of bias and variance in the distribution. They also state that this measure is mainly a univariate measure like that of Gelman and Rubin's convergence diagnostics. They state that if θ^t is taken as -2 times the log of the posterior density, one may be able to use it to investigate the posterior distribution. Cowles and Carlin (1996) state the disadvantages of the Geweke measure as its sensitivity to the specification of the spectral window and it is mainly based on the econometrician's experience and subjective view of the convergence diagnostic application.

Heiderberger-Welch Convergence Diagnostic:

Heidelberger and Welch (1983) combined the methods of Schruben (1982), and, Schruben et al. (1983) for detecting nonstationarity using the spectral analysis for estimating the variance of the sample mean. The Schruben (1982) and Schruben et al. (1983) methods are hypotheses test based on the application of statistics to the Brownian Bridge theory (see Hsu 1999) with an earlier run by length control procedure described in Heiderberger and Welch (1981a,b) (*explanation continues in the appendix*).

Heidelberger and Welch (1983) found that the stationarity test had little power to detect an initial transient when the run length is shorter than the extent of the initial transient, that is, when the whole sequence is in transient phase. This assertion is also corroborated by Brooks and Gelman (1998).

Ritter and Tanner (1992):

This involves a method they call the *Gibbs Stopper*. In this method an importance weight w is assigned to each $d-$dimensional vector, X, drawn at each Gibbs draw and then histograms are then drawn of each Gibbs draw obtained across multiple chains or across single chains. Convergence is achieved when the values appear tightly close together on the histogram. Wei and Tanner (1990) state that it is good practice to also look at the standard deviation values. They denote the appropriate weight as:

$$w = \frac{q(X)}{g_i(X)} \tag{6.5}$$

where $q(X)$ is proportional to the joint posterior (stationary) density and g_i is the current Gibbs sampler approximation to the joint posterior. The quantity q is always available in any Bayesian analysis because the joint posterior density is known to a normalising constant (*continues in the appendix*).

Brooks and Roberts (1998) state two problems with this diagnostic, first, it is time intensive and computationally expensive (Cowles and Carlin, 1996 also corroborates this point). They also state in the case of a single draws (replication), the weight estimated in

equation (6.5) might not tend to the correct value as $i \to \infty$ as it might just remain constant or "get stuck" in a particular part of the joint stationary distribution. This means a large chunk of the joint posterior density is left unexplored. They, thus, recommend the execution of multiple draws (replication). Another disadvantage of the method as stated by Cowles and Carlin (1996) is that the coding for the diagnostic is problem specific.

Zellner and Min (1995):

Zellner and Min in 1995 introduced a variant of the Ritter and Tanner (1992) method. The aim of their method is to determine convergence of the Gibbs sampler in a distribution and also to explore if it converges to right distribution. They proposed three *Gibbs Sampler Convergence Criteri*a. These are the "Gibbs sampler difference convergence criterion", the "anchored ratio convergence criterion", and, the "ratio convergence criterion". They state that their method is only applicable when parameters can be divided into two part α and β and when the joint posterior density function can be written down analytically.

Using the marginal density estimates of $\hat{p}(\alpha)$ and $\hat{p}(\beta)$ obtained by "Rao-Blackwellization" (please see Gelfand and Smith 1990 for an explanation of this term) for the parameters α_1 and β_1, we write the Gibbs sampler difference convergence criterion weight function as:

$$\hat{p}(\alpha)p(\beta|\alpha) - \hat{p}(\beta)(p(\alpha|\beta)) = \hat{\eta} \tag{6.6}$$

Zellner and Min (1995) state the if value of $\hat{\eta}$ is small in absolute terms and close to zero, then it can be said that convergence has occurred. Brooks and Robert (1998) state that this is similar to the weight function of Ritter and Tanner (1992).

The major disadvantage of this diagnostic as with Ritter and Tanner (1992) is that the explicit conditional posterior densities must be available. Also, Cowles and Carlin (1996) state that analytical work is needed when factorisation into two sets of parameters is not readily available. Like with the Ritter and Tanner (1992) method, this method is problem specific. The literature is not explicitly clear as to whether this diagnostic method may be extended to other Markov chain Monte Carlo methods (see Cowles and Carlin, 1996, p.888 and Brooks and Roberts 1998).

Normed Distance Criteria

As stated in Brooks and Roberts (1998), all of the methods we discuss in this section make use of the transition kernel of the sampler to monitor convergence of the target posterior distribution.

Liu, Liu, and, Rubin Convergence Diagnostic:

Liu et al. (1993) method employs a single statistic which they referred to as a "global control variable" to monitor the convergence of a joint posterior distribution. In other words, this variable is used to assess if a multi-dimensional Gibbs sampler has converged

to the stationary distribution. It is a more comprehensive diagnostic and it is not based on subsampling of the distribution. Cowles and Carlin (1996) describe the method thus; they state that the method requires running m parallel iterations started from an initial dispersed values, at iteration t, and for each independent parallel chains, i and j and the following statistic is constructed:

$$U_{ij}^{(t)} = \frac{\pi(X_j^t)}{\pi(X_i^t)} \frac{K(X_j^{t-1}, X_i^t)}{K(X_j^{t-1}, X_j^t)} \tag{6.7}$$

and $\quad U^t = \frac{1}{m(m-1)} \sum_{i \neq j} U_{ij}^t$

where $i \neq j; \quad i, j = 1, ..., m$ and X_j^t represents the vector of parameters generated at iteration t of chain j and $K(\boldsymbol{X}, \boldsymbol{Y})$ denote the probability of moving from X to Y in one iteration of the Gibbs sampler. Liu et al. (1993) propose two ways of using the "global control variable". First, is to divide the chains into $m/2$ distinct pairs and then construct the $\{U^t\}$ sequence for each pair. They propose the method of Gelman and Rubin (1992). The second is to generate the aggregate value of sample cumulative distribution of the function U and then use the Smirnov-statistic based tests to monitor convergence as recommended by Heidelberger and Welch (1983).

Brooks and Roberts (1998), and, Cowles and Carlin (1996) state that the major drawback of the convergence diagnostic test is that the variance of the global control variable may be very large. They both propose the use of the logarithm counterpart of U, unfortunately, the trade-off in this, is that it loses its direct interpretability which is an advantage of this method. They propose a second resolution of the problem; which is making the value of m very large, but they state that this might be computationally time intensive, however, with the advent of very fast modern computers this problem might be resolved easily.

Another disadvantage of this method is that it is problem specific as it requires new code to be written for each problem.

Roberts:

Roberts (1992,1996) intend to assess the convergence of the whole joint distribution. As stated in Cowles and Carlin (1996), Roberts method is applicable when the distribution of the iterates is continuous. His method only applies to Gibbs sampling. The method uses the norm distance between the marginal density of the symmetrised chain at θ^t (which is the density generated at the tth iteration) denoted by f_t and the (target) stationary density, f. This is denoted as: $\|f_t - f\|$ with $\|\cdot\|$ being the norm. The method uses the reversible chain Gibbs sampler which leads to a *dual chain* $(\tilde{\theta}^{(t)})$ (*explanation continues in the appendix*).

Robert and Casella (2004) lists three major drawbacks on the practicality of this diagnostic even though it is theoretically appealing. They assert that:

- because of the requirement of several parallel chains of iteration, it results in loss of efficiency in its application.

- the focus of the scalar control is based on f which is the marginal distribution of $\theta^{(t)}$, and of less importance in Marko chain Monte Carlo analysis

- the diagnostic can be time consuming, especially in the calculation of the normalising constant K and achieving stationarity around a mean value does not imply that whole chain has been properly and efficiently explored.

Brooks and Roberts (1998) state that other drawbacks include the fact that the variance can be very large, as well as, interpretation problems.

Yu L^1 Diagnostic:

Yu (1995) convergence diagnostic is based on the estimated L^1 error, such that the estimated L^1 error statistic captures any difference between the Markov chain density and the unnormalized target density when the Gibbs chain does NOT mix quickly or when it is sticky. The diagnostic is based on a single replication of the Gibbs chain. Cowles and Carlin (1996) state that the method assumes geometric ergodicity of the Gibbs sampler (*see appendix for explanation*).

Brooks and Roberts (1998) state that the diagnostic is problem specific and involves a lot computational drudgery. Also, the choice of the threshold point, 0.3, is rather arbitrary with no scientific proof for its choice. It also suffers from interpretation problems.

Yu also states that the diagnostic can suffer from false convergence if both the sampler and the region A miss the same major mode of $\pi(x)$. Brooks and Roberts (1998) also state that unlike Liu et al. (1993) and Roberts (1992) methods; the Yu (1995)'s method stabilises sample-wise rather than in expectation.

The Total Variation Diagnostic:

Brooks et al. (1997) introduce another convergence diagnostic that attempts to overestimate the total variation distance (L^1) between the full kernel estimate of the distribution from different replications. Brooks et al. (1997) started off by showing how the distance between two different probability distributions can be bounded from above using the rejection sampling idea as expounded in Smith and Gelfand (1992) (*see appendix for explanation*).

Brooks et al. (1997) highlight the interpretability and stability of their method and emphasize this as the advantage of their method over such methods like Roberts (1992), Liu et al. (1993). They also show how their method fits into different multivariate conditional densities.

Regeneration and Coupling Methods

The Johnson Diagnostics:

The Johnson diagnostic uses the method of the *coupled* Gibbs sampler. Brooks and Roberts (1998) explain the process of coupling time as the generation of two discrete time processes, X_1 and X_2 which are marginally distributed from the Markov chain in such a way that they are independent of one another, except that their joint distribution is Markovian (in practice, Brooks and Roberts 1998 state that they are actually constructed dependent of another). This process is referred to as coupling of X_1 and X_2. The coupling time τ is the process of equating the two discrete processes which is defined as :

$$\tau = inf\{t: X_i^t = X_2^t\} \tag{6.8}$$

Thus, if θ_i^0 is taken to be the stationary distribution π, they state the *coupling inequality* as:

$$\sup_A(\mathbb{P}[X_2^t \in A] - \pi(A)) \leq \mathbb{P}[][\tau > t] \tag{6.9}$$

Brooks and Roberts state the criteria for convergence to the target distribution as when the value of t is high (*explanation continues in the appendix*).

Brooks and Roberts (1998) state that the method is easy to apply and interpret. Johnson (1996) states a major drawback of the method is finding the appropriate 'overdispersed' estimate of the posterior distribution. They recommend running the procedure above several times in order to arrive at a suitable number of values for the output analysis.

The Propp and Wilson construction:

Propp and Wilson (1996) introduce another method of coupling. The method involves running a fixed time simulation backwards. As a result, they start by running the chain from a time -1 to zero but n times. They define a time t for all times as $-M \leq t < -1$. Brooks and Roberts (1998) give a summary of the process by assuming that chains are given as $X_1, ..., X_{|i|}$, where i is the Markov chain constructed in an ergodic way such that for $M \leq s < t \leq 0$

$$X_i^s = X_j^s \implies X_i^t = X_j^t$$

Coupling (convergence) is expected to occur at or before time 0. Propp and Wilson (1996) states that when M is chosen to be large then the values of t that occur during the backward simulation is most likely to be smaller than M. This method is rooted in 'backward' simulation.

Compared to the Johnson method, the Propp and Wilson method is more complicated and more difficult to implement. Brooks and Roberts (1998) state that this method is only limited to a small class of problems where the researcher can construct two state space processes simultaneously that is, "top" and "bottom". They believe the process is actually a construction and not actually a convergence diagnostic that has a promising future.

A.L.Gibbs Method:

Gibbs (2000) attempts to use the coupling method to determine the total number of iterations needed for convergence to take place.

Suppose X_t^1 and X_t^2 are two Markov chains on the same state space with same transition probabilities and with intial values x^1 and x^2 respectively. Coupling takes place at T^{x^1,x^2} if:

$$T^{x^1,x^2} = \min\{t : X_t^1 = X_t^2 | X_0^1 = x^1, X_0^2 = x^2\} \tag{6.10}$$

The mean coupling time is given as:

$$T = \max_{x^1,x^2} E(T^{x^1,x^2}) \tag{6.11}$$

where the maximum is taken over all possible initial states x^1 for X_t^1 and x^2 for X_t^2.

This method is not concerned with estimating the variance of estimates of expectations. This method has not been applied to situations where there is *full order* on the state space.

Mykland, Tierney and Yu (1995) method:

This is a regenerative simulation method. The regenerative simulation method is based on the fact that a stochastic process say $X_n : n = 0, 1...$ becomes regenerative if there are times $T_0 \leq T_1 \leq ...$ such that each T_i, is independent of the previous value and it is identically distributed. Then, one can say the before and after time values of the process is independent and identically distributed. These times themselves form a regenerative process; the process is delayed if $T_0 \neq 0$.

This is the principle behind Mykland et al. (1995) method. They suppose that if the renewal process has an equilibrium distribution π and the aim is to estimate $\theta = E_\pi[f]$ for some function f. If the process is ergodic, and, If the process is assumed to be finite with n number of "tours" then f will converge to θ. Also, if we let $N_i = T_i - T_{i-1}$ for $i = 1, ..., n$ and

$$Yi = \sum_{j=T_{j-1}+1}^{T_i} f(X_j) \tag{6.12}$$

The pairs (N_i, Y_i) are independent and identically distributed if $E[|Y_i|] < \infty$, and, $\hat{\theta} = \Sigma Y_i / \Sigma N_i = \bar{Y}/\bar{N} \to \theta$ by the law of large numbers and the *Renewal Theorem*. If the Y_i and N_i have finite variances, the distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$ converges to a normal distribution. In order to assess convergence, Mykland et al (1995) suggest plotting T_i/T_n against i/n which they referred to as the *Scaled Regeneration Quantile* (SRQ). If the renewal process has reached equilibrium, it is expected that the plot should almost be a straight line through the origin with slope equal to 1. They further state that any deviation from this, means that the chain needs to be run for longer and this is mainly due to the fact that some tours are longer others. They suggest an examination of the states traversed by the process during the longer "tours" which might suggest ways of improvement.

A drawback of the process is that it is problem specific. Also, the method the regeneration points are introduced into the Markov sampler is a problem, especially in continuous state space, Mykland et al. (1995) used the Nummelin (2004) method to introduce the regeneration points into the sampler in a general state space.

Robert (1995) Method:

In their method, they show that:

$$\frac{T_n}{n} \overset{n \to \infty}{\Longrightarrow} \mathbb{E}_\pi[T_2 - T_1] = \mu_A \tag{6.13}$$

and by virtue of the central limit theorem we have:

$$\frac{1}{\sqrt{n}} \sum_{n_0=1}^{n} (S_n - \lambda_{n_0} \mathbb{E}_\pi[h(x)]) \to N(0, \sigma_A^2) \tag{6.14}$$

with σ_A^2 estimated by

$$\hat{\sigma}_A^2 = \frac{1}{\sqrt{n}} \sum_{n_0=1}^{n} (S_n - \lambda_{n_0} \sum_{n_l=1}^{n} \frac{S_{n_l}}{N}) \tag{6.15}$$

As $\hat{\sigma}_A^2$ converges to σ_A^2, N/n converges to μ_A, thus $\dfrac{\hat{\sigma}_A n}{N}$ converges to σ_f^2

He states two drawbacks of this procedure; first, it requires the slowest ratio to converge for the algorithm to stop. Second, it is that the process is highly dependent on the choice of the starting values, since the closer the value of A, the faster the process will come to termination. This brings forth the question of how to define "close" in practice. Also, Brooks and Roberts (1996) state that the process is problem-specific and it is of limited practical use in the general state space. However, Robert (1995) states that this is probably the most promising method of assessing convergence.

The Eigenvalue Bounds Methods

Garren and Smith (2000) attempt to show that convergence occurs when the second largest eigenvalue of the Markov transition matrix is close to one. Assuming the objective is to estimate $\rho = \Pi(D)$, they define $Z_n = I(X_n \in D)$ where $I(\cdot)$ is an indicator function and D is a non-empty proper sub-set such that $\rho_n = E(Z_n)$. They also assumed that ρ is between zero and one.

In order to estimate ρ, they assume that the transition operator is self-adjoint and of the class *Hilbert-Schmidt* giving:

$$\rho^n = \rho + a_2 \lambda_2^n + O(|\lambda_k|^n) \tag{6.16}$$

whereas $n \to \infty$, *for some* $k \geq 3$ *and* a_2 *is some real number* and λ_k is eigenvalue from $1, ..k$ such that $|\lambda_2| > |\lambda_k| = 1$. Chan and Geyer (1994) state that the strict inequality among the eigenvalues does not need to occur, especially in the Metropolis-Hastings sampling where candidate distribution are not accepted with probability 1 (*see appendix for further exposition*).

Cowles and Carlin (1996) list the major disadvantage of this method as its dependence on a large number of replications which must all share the same starting point, as a result, they do not seek to explore different regions of the state space. They also state that this method only attempted to solve burn-in problems without resolving the problem of variance in estimation, while asserting that the plots are very difficult to interpret and involves a lot of computational drudgery.

Raftery and Lewis:

Raftery et al. (1992) while considering the number of iterations needed for convergence to the stationary distribution to occur, also focuses on estimating quantiles of functional of the posterior distribution.

Suppose the problem is to estimate the particular quantiles of the posterior distribution of a function U with parameter θ, thus we want to estimate $P[U \leq u|y|]$ to within $\pm r$ with probability p where U is a function of θ. They propose a method to determine the following; the length of the burn-in periods, M, the amount of further N iterations that the algorithm needs to run for, and the storage of every k^{th}. In summary, the method is all about the determination of M, N, and k. The method uses a two-stage Markov process to determine these three quantities.

Brooks and Roberts (1998) state that this method only provides convergence values for the quantile and not for the whole distribution. They state that most software tends to use the Raftery and Lewis value for q and r but state that this may lead to strong underestimation of the true length of the burn-in period and should therefore not be accepted as a general

rule. Maceachern and Berliner (1994) prove that discarding samples is not a good idea as it often leads to inefficient estimation.

Time Series Methods

These methods see the Markov chain as asymptotic time series.

The *Phase Randomisation* Method:

In this method, Nur et al. (2005) see the Markov chain Monte Carlo simulation as a time series. This method involves taking the Fourier transform of a given time series and replacing the *phase* with a value sampled uniformly on $(0, 2\pi)$ and back transforming to render a *surrogate* series.

Thus, they state that suppose X_t is a time-series and $X = (X_1, ..., X_N)$ is the dataset and if $E(X_t) = \mu$, $\gamma_k = E((X_t - \mu)(X_t + \mu))$ are the expectation and autocovariances of X_t. A fictitious data $Y = (Y_1, ..., Y_N)$ is generated by the use of the *surrogate* data such that $N = 2m + 1$, and,

$$Y_t = \bar{X} + \sqrt{\frac{2\pi}{N}} \sum_{j=1}^{m} \sqrt[2]{I(X, \omega_j) \cos(\omega_t j + \theta_j)} \qquad for \ t = 1, ..., N \qquad (6.17)$$

where I denotes the modified periodogram and ω_j denotes the angular frequencies such that $\omega_j = 2\pi j/N$ and $\theta_1, ..., \theta_m$ are independent and identically distributed uniform random numbers with mean, 0 and variance, 2π. This process is continued until a *surrogate* series is achieved. They use the *Rescaling Surrogate* algorithm of Theiler et al. (1992) to generate a bootstrap *surrogate* series.

They summarise the steps for assessing the convergence as follows:

1. Generate an original Markov chain using a chosen Markov sampler like Gibbs

2. Then, generate the *surrogates* of the original chain by applying the rescaling algorithm

3. The third cumulant estimates of the surrogates in step 2 above are then plotted.

Convergence occurs if the plotted histogram is unimodal around zero notwithstanding the variance value as long as the non-zero histogram does not have a long tail or it is multimodal of which the chain is said not to have converged yet. They further advised that the Kolmogrov-Smirnov or Shapiro-Wilk tests for normality is adopted in further confirming convergence.

Nur et al. (2005) state that their method could help in identifying appropriate subsamples for the Metropolis-Hastings algorithm. Also, they believe that the multiple chain version

of their method is possible. They, however, state that their method has a major draw-back of being computationally burdensome. Our view is that the method is likely to be inefficient has it depends so much on the quality of its subsamples.

Distance Methods

These methods measure the distance between Markov chains to assess convergence.

The Hellinger Distance Method:

Boone et al. (2014) introduced this method. The measure focuses on measuring the *Hellinger* distance between two Markov chains with the same stationary distribution (*see appendix for further elucidation*).

They state that one major drawback of their method (as with kernel density estimation) is to make sure that the variance ,σ^2, is not close to zero. This is called the *edge effect.* If the distribution of the variance is close to zero, then, depending on the kernel used, then the density estimate for the variance, σ^2, may give positive or negative probability values, which gives inaccurate results of the diagnostic.

Brooks, Giudici and Philippe Methods:

Brooks et al. (2003) use non-parametric model selection methods to assess convergence. They use several distance measures to assess how close the replications of the chains are to the stationary distribution. Convergence occurs when the distance between replications are small.

Consider J independent Markov chains for a total T iterations. For each chain $j = 1, ..., J$ they estimate a probability mass function denoted by $\hat{P}^j = (P_1^j, ..., P_c^j)$ for ease of notation, T, is dropped. If T is sufficiently large, then convergence occurs when the estimates of the probability mass function for each of the chains are similar. They suggested running the chain for $2T$ iterations with the first T being the burn-in. The performance of the J chain is assessed as the distance between the mass functions approaches zero.

Though, their method is Markov based they state that assumption of approximate independence is important for their method to work. They, thus, showed how this could be done. They recommend subsampling each chain at each λ observation (λ is a thinning parameter which can be tuned) as one way creating independence between chain. They also recommend the use of the tests like the Kolmogorov-Smirnov and the chi-squared statistics to achieve this aim.

This method suffers from interpretation problems. Different models might require different thinning factors. This, for us, makes this method problem specific.

The Batch/Stratification Methods

This method divides the chains into batches (subsamples) and then assess convergence using these batches.

The Subsampling Method:

Giakoumatos et al. (1999) develop a method of using subsampling to assess convergence. They use the subsampling methodology of Politis and Romano (1994) and Politis et al. (1997). They demonstrated their method for both simple and multiple Markov chains. They did not pursue the issue of confidence intervals approximately "coinciding" in each Markov chain (*see appendix for further exposition*).

The Stratification Method:

In this method, Paul et al. (2012) use the stratification and post-stratification methods of Jones et al. (2006) and Flegal et al. (2008) to develop a way of combining different subchains of different posterior expectations of Markov chains. They then develop variance estimates for the Markov chain-based estimators. Convergence assessment is then based on the variance estimates. The approach also attempts to assess the level of mixing of a Markov chain (see appendix for further exposition).

A major drawback is that there is no specified method in their work that shows how the strata will be selected. Since, the method is based on assessing the convergence of batches, it does not tell us if the chain, as a whole, as converged.

The PACE Method:

VanDerwerken (2015) in his unpublished PhD thesis introduced the *Partition-based Approximation for Convergence Evaluation* (PACE) into the literature on convergence diagnostics. The method involves generating multiple chains at overdispersed locations in the state space. Replications (Draws) are collected at different location in the J chain and clustered to obtain partitions. The distance between the sample distribution each individual chain are then calculated. The author state that the distance calculated will involve a within-chain and across-chain probabilities of the partitions. Convergence is attained at the point where the percentage of within-chain draws belonging to a given element C will be approximately equal to the across-chain draw values. The value obtained through this procedure is the maximum estimated approximate L^1 across J chain. The author state that their method helps to solve the issue of dimensionality in the literature. The major drawback of the method is that it assesses convergence of the partition not the whole distribution.

Philippe and Robert Riemann Sums Method

Philippe and Robert (2001) used the Rao-Blackwellised version of the Riemann sums to assess convergence to the stationary distribution of unidimensional and multidimensional chains (see appendix for further exposition).

This method uses an approximation to the Riemann sum which means it is likely that the iteration only takes place in a particular part of the distribution, in other words, the chain could be sticky. They did not explicitly state how to monitor or resolve this issue when it arises. We believe that the method is likely to involve a lot of computational time also.

The Fixed Width Approach:

In his PhD thesis, Gong (2015) uses the *fixed width stopping rule* procedure of Jones et al. (2006) to monitor the convergence of the chain to a stationary distribution. This is a confidence interval based procedure, where the simulation is stopped at the point when the width of a confidence interval is close enough in relation to the size of the stationary distribution. He introduces two stopping rule which are in relation to the stationary distribution - these are the magnitude and standard deviation of the fixed-width stopping rule. The procedure is dependent on the functional central limit theory and a strong estiamtor of the asymptotic variance such that $\hat{\sigma}_n^2 \to \hat{\sigma}_\theta^2$ as $n \to \infty$ where $\theta \, \epsilon \, \mathbb{R}^p, \;\; p \geq 1$.

Suppose we have a Markov chain, the fixed-width rule states that one can construct a $(1 - \delta)100\%$ confidence interval for θ with width:

$$\omega_\delta = {}^{2}z_{\delta/2} \frac{\hat{\sigma}_n}{\sqrt{n}} \tag{6.18}$$

where ${}^{z_\delta}/_2$ is a critical value from the standard normal distribution and $\hat{\sigma}_n^2$ is a strongly consistent estimator of σ_n^2.

This method is based on finding a strongly consistent estimator for the variance which might not be an easy task. We believe the procedure is promising but we do not know how much computatinal drudgery is involved as the procedure has not been tested by other researchers. The author further affirmed on page 55 of his work that Bayesians are still being challenged on finding the best way to determine how well the chain mixes and explores the state space.

At this juncture, we are confident to present the summary of this chapter.

6.5. Summary

We attempted to review recent literature on Markov chain Monte Carlo convergence with special focus on the use of diagnostics. Our research did not focus on the theoretical method of convergence assessment. We discussed convergence diagnostics under the graphical and formal approach headings. One could also classify them under the Single Chain Output Diagnostics and the Multiple Chain Output Diagnostics.

The single chain output diagnostics include the methods under Spectral Analysis methods (Geweke et al. (1991) and Heidelberger and Welch (1983) methods), the Yu and Mykland (1998), Zellner and Min (1995), Raftery et al. (1992), and, Yu and Mykland (1998) methods.

The multiple chain output diagnostics include the methods introduced by the following authors: Liu et al. (1993), Ritter and Tanner (1992), Garren and Smith (2000), Johnson (1996), Mykland et al. (1995), Roberts (1992,1996), Brooks and Gelman (1998),Philippe

and Robert (2001), VanDerwerken (2015) and, Gong (2015). For a comparative review of most of the methods, we refer the reader to Brooks and Roberts (1998), Cowles and Carlin (1996), and, Robert and Casella(2004, Ch. 12).

However, we state that none of these methods are foolproof, also, the values obtained, as well as, its interpretation is greatly affected by the way each of the developers of these methods defines the word "convergence" (see Cowles and Carlin 1996, p. 899, Brooks and Roberts 1998 for explanation). This underlies our initial statement on the universally acceptable definition of "convergence". The literature recommends the use of several diagnostics in order to ascertain if the chain has converged. As it has also been noted that different diagnostics may give different results of which it is now down to the researcher's discretion to either continue the sampling or not.

For ease of comparison and to reduce the time of the researcher writing codes for his choice of convergence diagnostics, Best et al. (1995) develop a collection of S-plus routines in Bayesian Analysis Using Gibbs Sampling (BUGS) called CODA. The CODA function has been modified into the Matlab® version by Jim Lesage which can be found in his Matlab® toolbox, which we make use of in this thesis. We use the autocorrelation and Raftery-Lewis code of the CODA function. The choice of the two diagnostics is based on the preference the researcher and not on any empirical factors. The results are presented in chapter nine of this thesis. These diagnostics can also be used by frequentist econometricians.

7. Data Collection And Description

7.1. Introduction

In the last three chapters, we have reviewed the literature on the household model, the stochastic frontier model, and, the convergence diagnostics. In this chapter, we discuss the data set for this research - how they were collected, some idiosyncrasies associated with them, and problems faced in their collection. The data set is secondary in nature and is from three different African countries: Nigeria, Ethiopia and Tanzania (we gave a brief description of the chosen areas for our research in chapter two). Also, in this chapter, we discuss the malaria data set, which we obtained from the Malaria Atlas Project (MAP) in Oxford University. The importance of the malaria data set in this research cannot be over-emphasized. As a result, one of the main focus of this research is to describe how we synchronise each of the individual household data set from the different countries with the malaria data set. Apart from the availability of reliable household data from these three countries, one other reason for choosing these three different countries is the variation in their malaria prevalence value this we depict this in figures 7.1, 7.2, and, 7.3 below (the Malaria Atlas Project describes these countries as *Plasmodium falciparum* stable areas). We describe the household data set of the different countries with their corresponding descriptive statistics in sections 7.2, and discuss the spatial malaria data set in section 7.3 and the merger of these two data set in section 7.3.2 below.

Figure 7.1.:   Plasmodium falciparum Malaria Transmission Map In 2010 In Nigeria
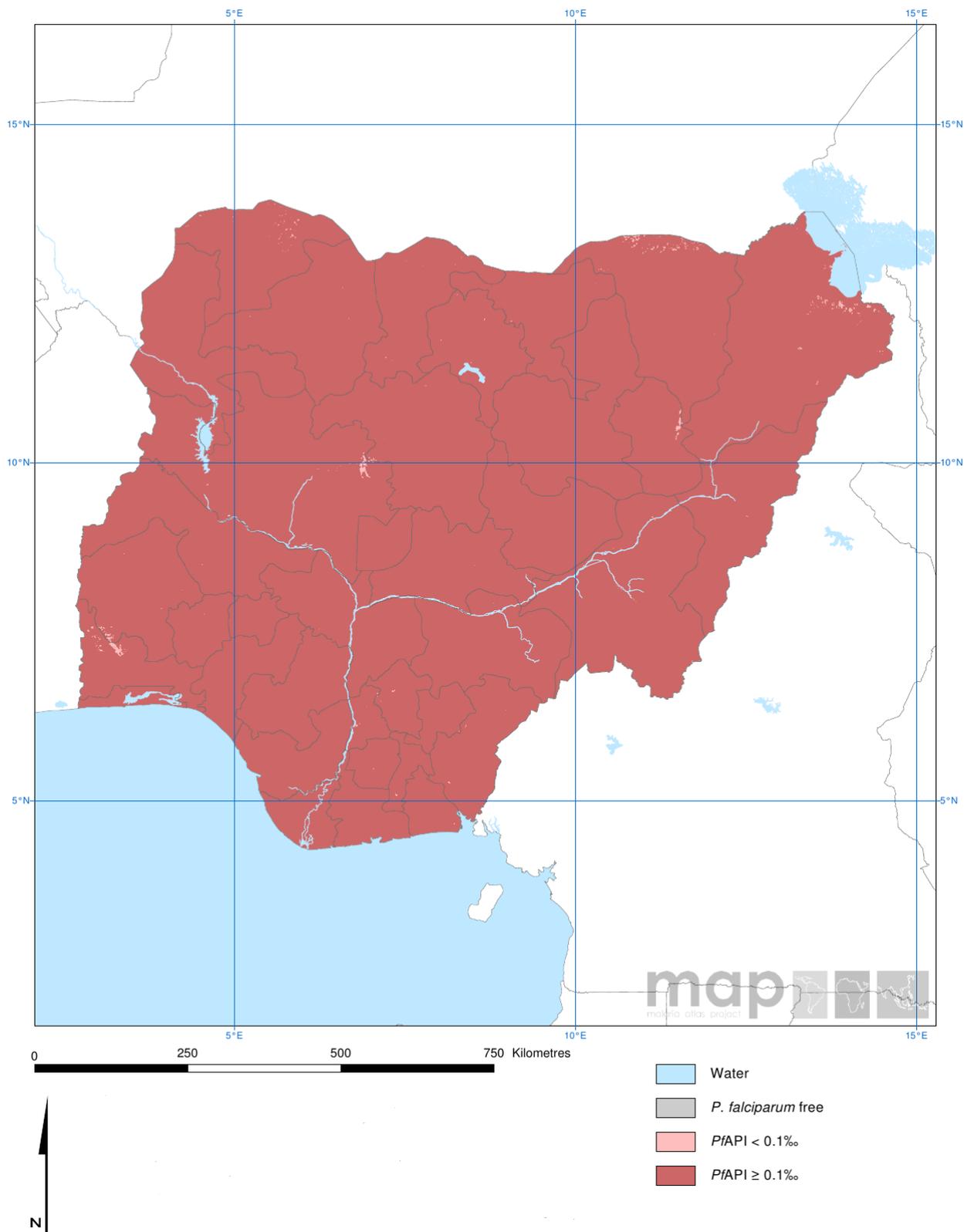
Pink areas are malaria unstable, and, Red areas are malaria stable. Nigeria has no grey coloured area and it is one of the four countries that accounts for half of the World's malaria burden.

Adapted from MAP website: www.map.ox.ac.uk. Retrieved Nov. 2013

Figure 7.2.:   Plasmodium falciparum Malaria Transmission Map In 2010 In Ethiopia
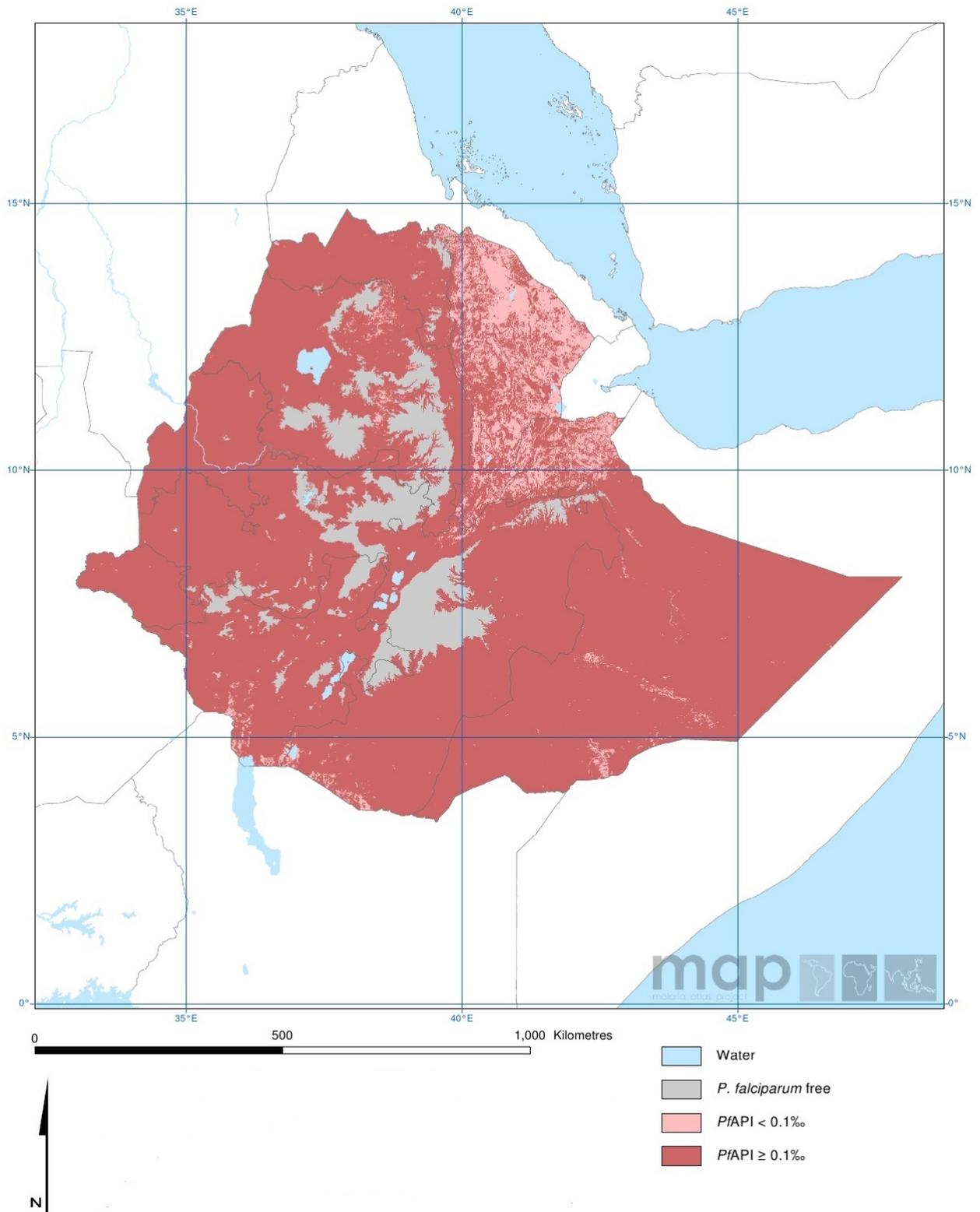
Grey areas are malaria free, Pink areas are malaria unstable, while, Red areas are malaria stable. Ethiopia has more areas coloured grey than the two other countries we consider in this research

Adapted from MAP website: www.map.ox.ac.uk. Retrieved Nov. 2013.

Figure 7.3.:   Plasmodium falciparum malaria Transmission map in 2010 in Tanzania (United Republic of)
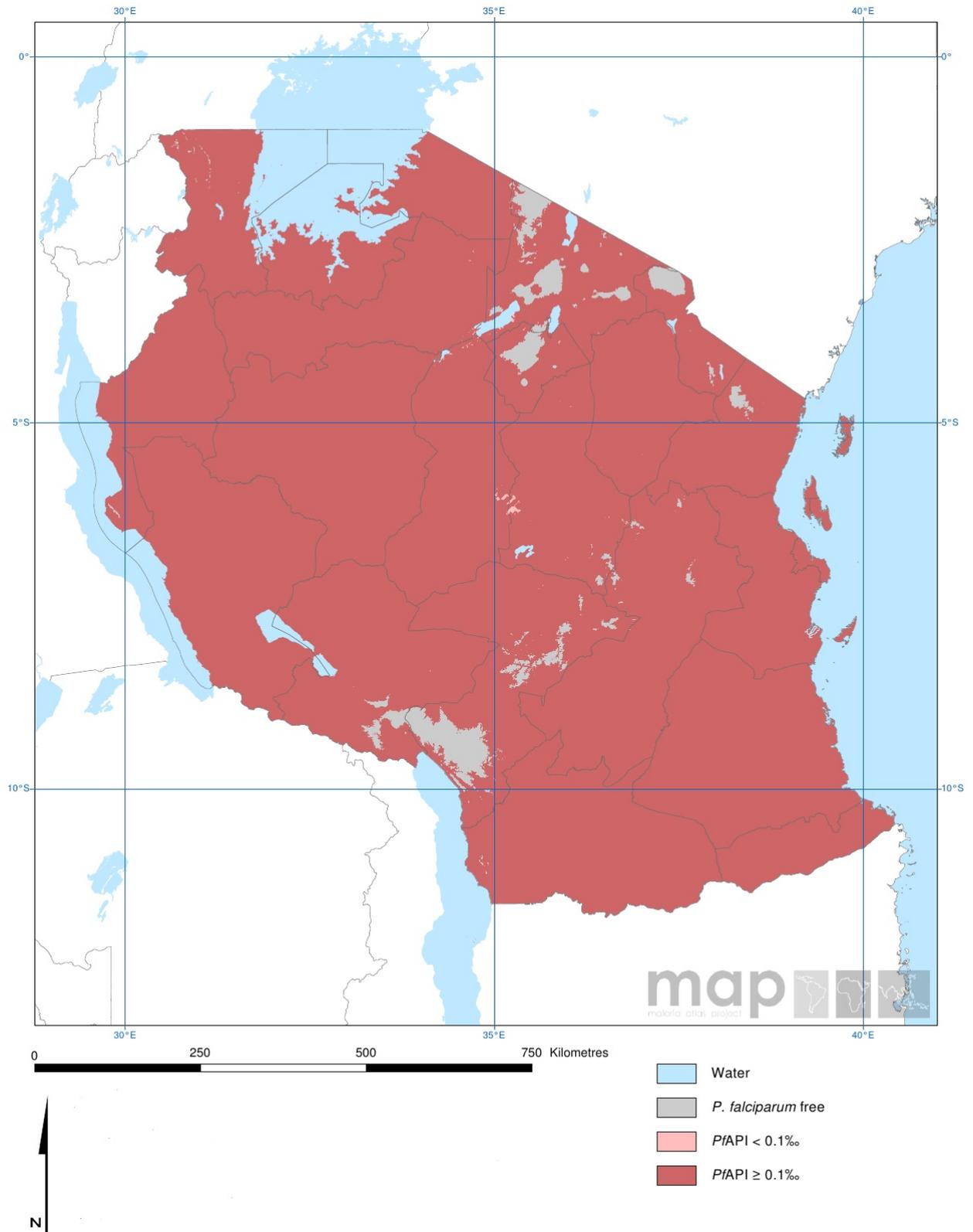
Grey areas are malaria free, Pink areas are malaria unstable, and, Red areas are malaria stable. Tanzania has more areas coloured grey than Nigeria.

Adapted from MAP website: www.map.ox.ac.uk. Retrieved Nov. 2013

7.2. The Data

7.2.1. The Nigeria Data Set

We obtain the Nigeria data from the Nigeria version of Living Standard Measurement Study project. This is a World Bank project, which initially started in the year 1980 with the sole aim of improving the quality of household data collected in developing countries for policy decision making. However, in early 2000, the Living Standard Measurement Survey was expanded to include an agricultural survey aspect, hence, the survey was recodified The Living Standard Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA). The sole aim is similar to the initial objective except that it now focuses on the production of reliable agricultural panel data for seven Sub-Sahara African countries. The Bill and Melinda Gates Foundation funds this project in collaboration with the World Bank and the individual countries statistical offices.

Though, the Nigerian National Bureau of Statistics (NNBS) carried out the Harmonised Living Standard Survey (HNLSS) in 2003/2004 but this survey lacks in-depth information on the Agricultural component as it relates to household welfare and poverty. We note that the samples used for this survey are different from that used for the Living Standard Measurement Survey - Integrated Survey on Agriculture. In carrying out, the Living Standard Measurement Survey - Integrated Surveys on Agriculture, the Nigerian National Bureau of Statistics revises one of its core surveys, the General Household Survey (GHS) and added a panel component to it and renamed the LSMS-ISA survey the General Household Panel Survey.

The General Household Panel Survey is a survey of twenty-two thousand households in Nigeria, from this number, a total of five thousand households are selected for the panel survey aspect of the project which also covers the agricultural aspect of the survey. The first phase of the project in the first round was divided into two stages - the post-planting and the post-harvest aspects of the survey. The post planting aspect of the survey was carried out in 2010, while the post harvest from carried out in 2011. The households selected for the panel aspect of the survey are re-interviewed two years after the time of the previous survey.

The NBS Nigeria (2011) documents lists some unique features of the survey as the creation of a panel data to study Nigeria's poverty dynamics as it relates to agriculture over time and space, use of the Geographical Positioning System to measure the size of agricultural land, and the adoption of the concurrent data entry method while collecting the data (these features are the same for all LSMS-ISA surveys in Sub-Sahara Africa). At the time of this research, the first two rounds (post-planting and post-harvest) have been completed and published.

Next, we discuss the selection of samples for the Nigeria General Household Survey.

Selection of Sample for the Panel Survey

Prior to carrying out the main survey itself, the Nigerian National Bureau of Statistics carried out a pilot survey. They select six states in this survey, these are Kaduna, Nassarawa, Taraba, Osun, Edo, and, Enugu states of Nigeria. They then select two enumeration areas (one urban and one rural) from each of these states. In each of the two selected enumeration areas, ten households were randomly selected, giving a total of sixty households used for the pilot survey (we note that enumeration areas are also defined as clusters in the literature).

For the main survey itself, the 2006 Housing and Population Census provide the *National Master Sample Frame* (this census delineates the whole country into 662, 000 enumeration areas consisting of 774 local government areas). Excluding the six Local Government Areas of Nigeria's Federal Capital Territory, Abuja, the Nigerian National Bureau of Statistics selects thirty enumeration areas from the remaining six hundred and sixty eight Local Government Areas. In each of the remaining six Local Government Areas of the Federal Capital Territory, they select forty enumeration areas; hence, twenty-three thousand, two hundred and eighty enumeration areas are selected, which then serves as the *Master Frame* for the survey.

According to Nigeria's NBS Nigeria (2011), from this *Master Frame* they develop another *Master Sample Frame* different from the earlier stated national master sample frame. This sample frame was developed by pooling the selected enumeration areas in the master frame by state. Then, a systematic sample of two hundred enumeration areas is selected from each individual state by ensuring that all the enumeration areas had equal chances of being selected.

These selected enumeration areas form what the Nigerian National Bureau of Statistics referred to as the National Integrated Survey of Households 2007/2012 Master Sample Frame. They, however, further emphasized that the sample for the LSMS-ISA survey is a sub-set of the National Integrated Survey of Households 2007/2012 Master Sample Frame. We discuss the selection of the sample from these master sample frame next.

Sample Frame for the LSMS-ISA Survey

Five hundred enumeration areas were interviewed for the panel aspect of this survey with ten households selected in each of the selected enumeration areas, giving a total of five thousand households for the survey. This process involves a two-stage stratified random sampling procedure. In the first stage, the enumeration areas were selected proportional to the size and total number of households listed in the enumeration area. Thereafter, ten households were selected from each enumeration area. In doing this, they first calculate the sampling interval by dividing the total households listed in each enumeration area by ten. A random number was selected from the table of random numbers and the household that this number corresponds to serves as the first selected household. Subsequent households were selected by the addition of the sampling interval to the preceding household

number until the 10th household was reached (*see the appendix for table on distribution of enumeration areas*).

The Data Set File

The Nigerian LSMS-ISA survey consist of three different questionnaires which are the household, agriculture, and, the community questionnaires. Each one of these differs in the type of questions it contains and we detail this in the appendix. The household file consists of sixteen data files, the agriculture file consists of thirteen data files, while, the community file consists of five data files. The files are named by adding the prefix *SECT* followed by the section number, for example section 1 of the household data is written as *SECT1.* It differs in a situation where you have sub-sections, in this case a number indicating the reference period follows the section number. The naming of the agricultural file also differs slightly because it includes alphabets after the section number and where you have sub-sections a number follows the alphabet.

Merging of the Different Sections of the Data

It was easy to merge the various sections (this data is cross-sectional in nature) together because the unique household identification number for our samples had already been constructed in the data. They generate the unique household identification number by concatenating the Zone, State, LGA, Sector, EA, RIC and HouseHold Identification variables.

The data collectors highlighted some problems encountered during the time of the collection of the data and these include the lack of electricity, malfunctioning of the Global Positioning System equipment, and, flooding.

7.2.2. Sample Description of the Nigeria Data

In this section, we describe the main features of our data, this is done in order to gain better insight and understanding of the socio-demographic characteristics of the Nigeria data used in our analysis. Thus, we present the results of the socioeconomic characteristics of the Nigeria data in this section.

Socioeconomic Characteristics of the Farmers

The socioeconomic description of the Nigeria sample reveals that the majority (32 percent) of our sample has a total output of one thousand kilograms and below (see table 7.1 and figure 7.4). About 17 percent have a total output of between one thousand kilograms per hectare and two thousand kilograms per hectare, while about thirty-nine percent of the farmers have an output of between two thousand and twenty thousand kilograms per hectare. No farmer had an output between eighteen thousand and nineteen thousand kilograms per hectare, while, about two percent of the farmers have output above twenty thousand kilograms per hectare. The farmers in the Nigeria sample produce an average of about four thousand three hundred kilograms per hectare of crops in total; the maximum total output is one hundred and five thousand kilograms per hectare. We discover that

most (over ninety five percent) farmers in the Nigeria sample are below fifty years of age, which implies our sample is made up of young farmers; this may influence their ability to take risks as younger farmers tend to be less risk averse than older farmers (see table 7.3 and figure 7.5 for further explanation).

All the farmers possess one form of education or the other (see table 7.4 and figure 7.6). About forty-four percent of the farmers had primary school education, while about thirty-eight percent of the farmers had secondary school education; and less than four percent have either teacher training education, vocational education, or National Certificate of Education. None of the farmers possess a higher degree certificate, about twelve percent possesses either quranic or adult education.

Most (about eighty-two percent) of the farmers operate on land size of less than two hectares in size, about seventeen percent of them have size of land of between three and ten hectares, while a little under one percent possesses land above ten hectares, hence, most of our respondents are small scale farmers (see table 7.5 and figure 7.7). This may impact on farm efficiency, Cornia (1985) observes that small scale farmers tend to produce more output per unit of land than large scale farmers. We observe that seventy-six percent of our respondents source for land through the family or community (see table 7.6 and figure 7.8). This may lead to further fragmentation of land and could be one of the reasons why farming is carried out on small land size in Nigeria. The second most (ten percent) popular way of having access to land in Nigeria is through its use free of charge, this could be as a result of the type of social interaction that exists among the members of the community in our sample area. Other less popular means of land ownership are, through outright purchase (eight percent) and rentage (six percent) respectively.

From table 7.7 and figure 7.9, mixed cropping remains the most popular cropping method in Nigeria, with sixty-eight percent of the farmers practising this method. Monocropping is the second most popular cropping method in Nigeria, with about twenty-five percent of farmers adopting this method. Intercropping and relay cropping are not popular among the farmers, with values of about four and two percent respectively. Alley cropping is the least most popular cropping method with less than one percent of the farmers practising this method.

Table 7.1.:   Distribution of Farmers by Total Output of Crops in the Sample

Output (kg/ha)	Frequency	Percentage (%)
0-1000	131	32.03
1001-2000	70	17.11
2001-3000	44	10.76
3001-4000	46	11.25
4001-5000	24	5.87
5001-6000	15	3.67
6001-7000	11	2.69
7001-8000	15	3.67
8001-9000	6	1.47
9001-10000	11	2.69
10001-11000	7	1.71
11001-12000	5	1.22
12001-13000	3	0.73
13001-14000	2	0.49
14001-15000	1	0.24
15001-16000	3	0.73
16001-17000	3	0.73
17001-18000	1	0.24
18001-19000	0	0.00
19001-20000	1	0.24
More	10	2.44
Total	**409**	**100%**

Figure 7.4.:   Histogram Showing Distribution of Farmers by Total Output of Crops
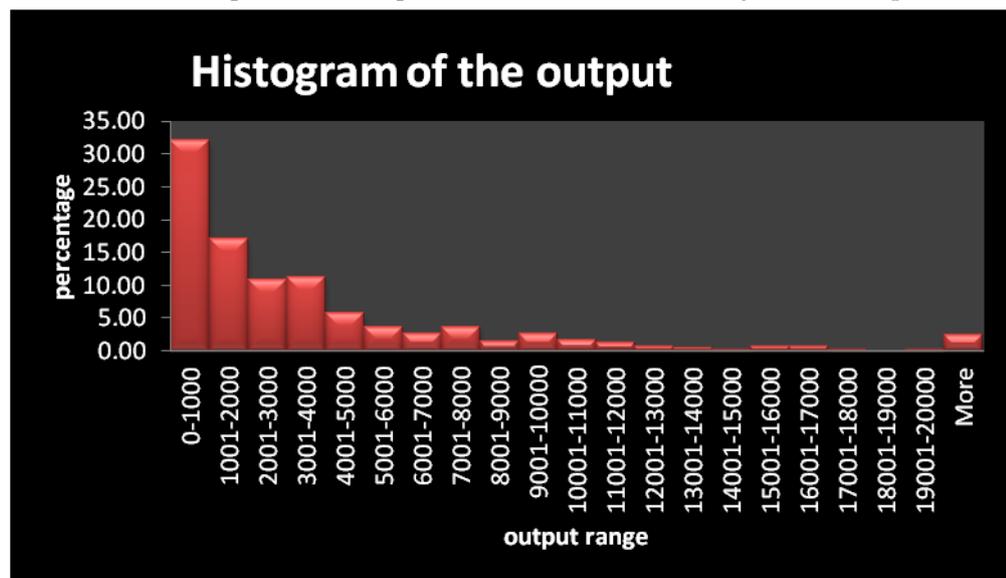
Figure 7.5.:   Histogram Showing Age Distribution of the Farmers
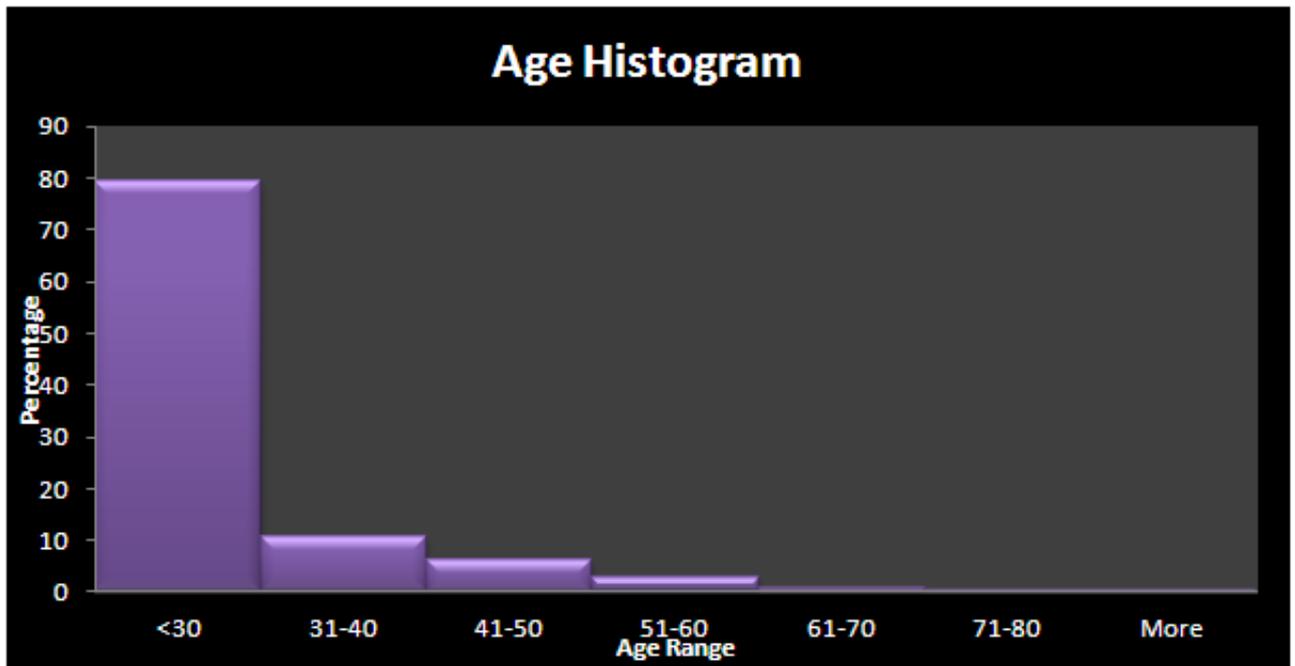
Table 7.3.:   Tabular Representation of Farmers' Age

Age (yrs)	Frequency	Percentage
<30	324	79.22
31-40	43	10.51
41-50	26	6.36
51-60	11	2.69
61-70	3	0.73
71-80	1	0.24
More	1	0.24
Total	**409**	**100.00**

Figure 7.6.:   Histogram Showing Education Distribution of the Farmers
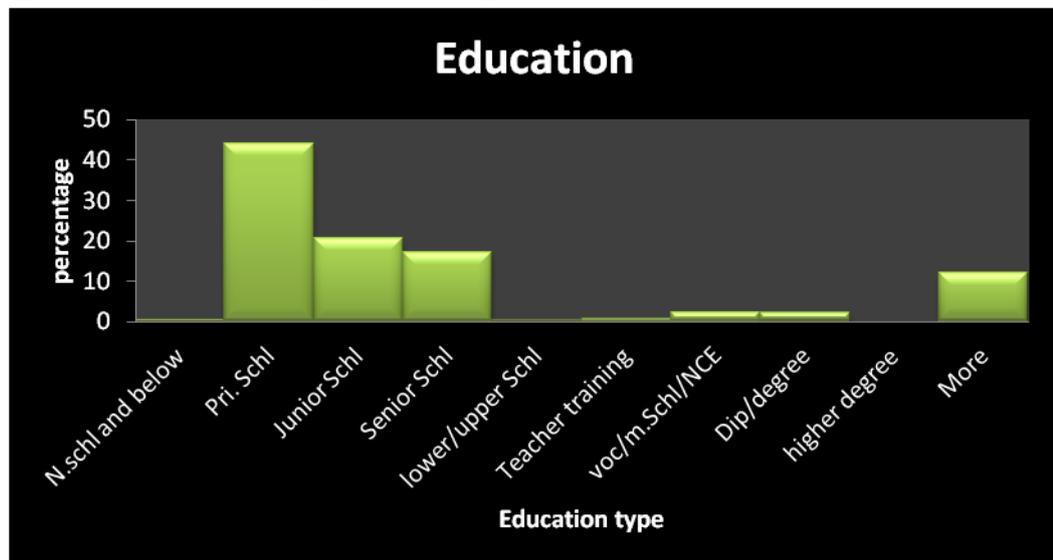
Table 7.4.:   Tabular Representation of the Farmers' Education

Education	*Percentage (%)*	*Frequency*
Nursry school and below	0.49	2
Primary School	44.01	180
Junior School	20.78	85
Senior School	17.11	70
Lower/Upper School	0.24	1
Teacher Training	0.73	3
Vocational Training/Modern School/NCE	2.44	10
Diploma/Degree	2.20	9
Higher Degree	0.00	0
Others	11.98	49
Total	*100*	*409*

Figure 7.7.:   Histogram Showing Land Size Distribution of the Respondents
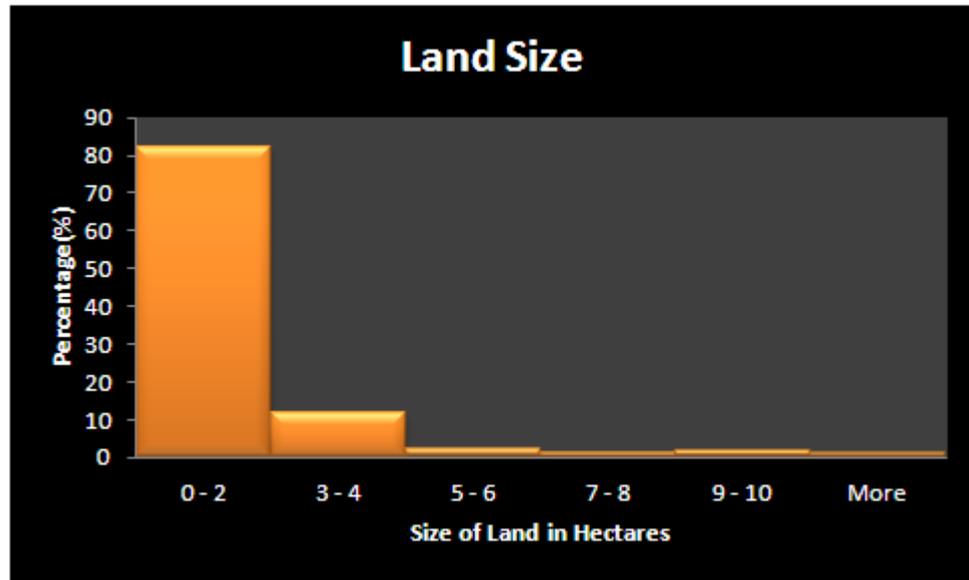
Table 7.5.:   Tabular Representation of the Land Size

Land (ha)	Percentage (%)	Frequency
0 - 2	81.91	335
3 - 4	11.98	49
5 - 6	2.20	9
7 - 8	1.22	5
9 - 10	1.71	7
More	0.98	4
Total	**100.00**	**409**

Figure 7.8.:   Histogram Showing Source of Land Ownership
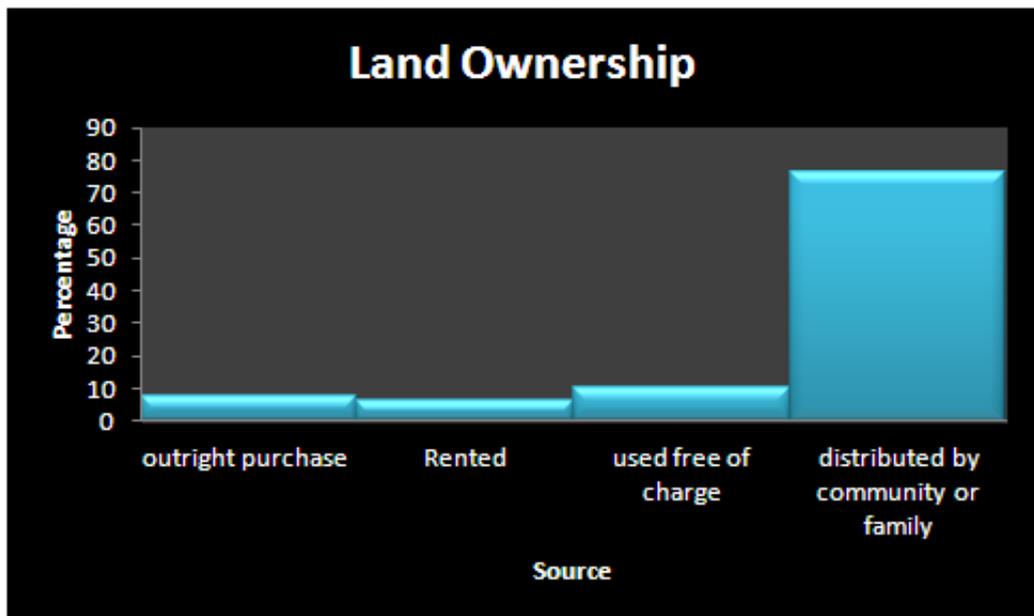
Table 7.6.:   Tabular Representation of Land Ownership

Source of Land Ownership	Percentage (%)	Frequency
outright purchase	8	31
Rented	6	25
used free of charge	10	41
distributed by community or family	76	312
Total	**100**	**409**

Table 7.7.:   Tabular Representation of the Cropping Methods

Cropping Methods	*Percentage (%)*	*Frequency*
monocropping	24.9	102
intercropping	4.4	18
relay-cropping	2.2	9
mixed cropping	68.0	278
alley cropping	0.5	2
Total	*100.0*	*409*

Figure 7.9.:   Histogram Showing Cropping Method
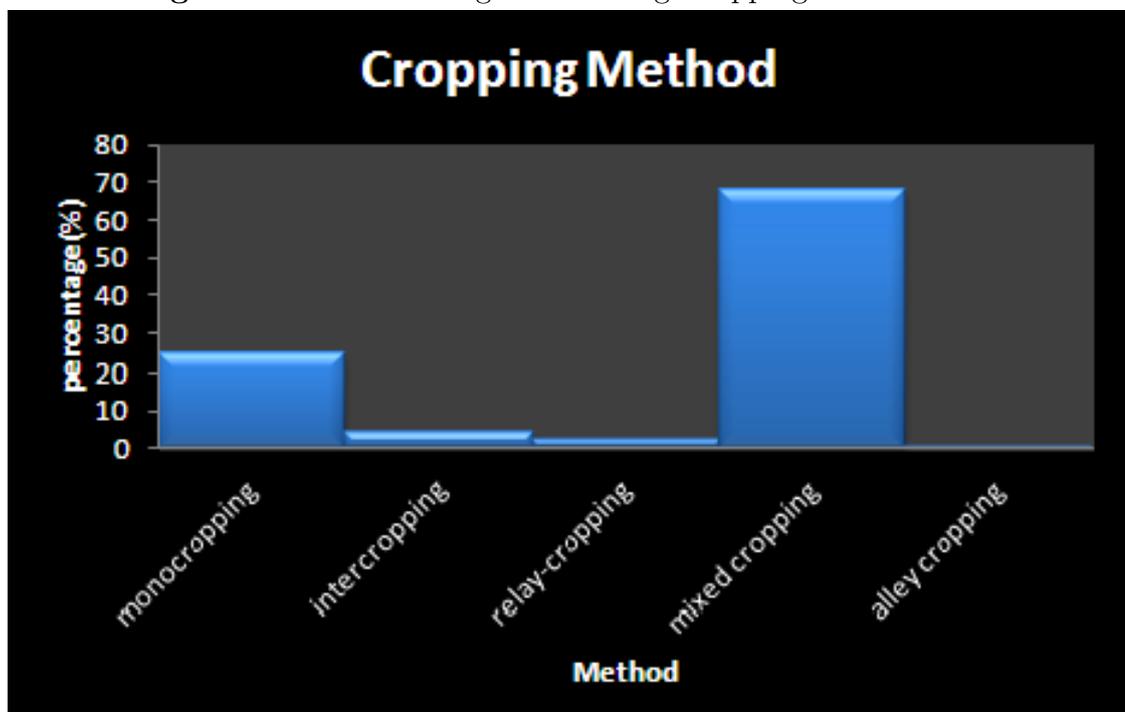
7.2.3. Characteristis of the Nigeria Variables

An important procedure in arriving at an empirical result is the definition of the variables to be used in the analysis. The variables we include in this analysis are based on those identified in the literature and in the survey documents of Nigeria. We present a general descriptive statistics of all the variables we employ in this analysis in table 7.8 below. It is hoped that this will help the reader to further understand the nature of our data set and also to properly assess the distribution of the each of the variables.

Table 7.8.:   Descriptive Statistics of the Sociodemographic Variables in the Nigeria Data

Variables	Description	Min value	Mean	Median	Max	Std Dev
Irrigation	Is plot irrigated. yes=1, No=2	1	1.96	2	2	0.21
land_size	Total land cultivated in sq. metres	58	16,748.64	7, 004.60	768, 539.40	50, 610.98
Hrs_HH	Total hours household labour use on the farm per season	0	14.70	10	172	16.59
Hrs_Hdlab	Total hours hired labour uses on the farm per season	0	173.16	170	2910	401.45
Seed	Total amount of seed used in production (kilograms)	0	306.44	15	25,736	2,107.10
Land_owner-ship	Type of land ownership, 1=outright purchase, 2=rented for cash or inkind,3=used free of charge, 4=distributed by community of family	1	3.54	4	4	0.90
Pesticide	Did the farmer use pesticide or not? Yes=1, No=2	1	1.76	2	2	0.43
machinery use	Did the farmer use any machinery or machine on plot? Yes=1, No=2	1	1.73	2	2	0.44
fertiliser	Laspeyres type of total fertiliser used in kilograms	0.5	69.65	50	499.50	84.20
Herbicide	Did you use herbicide on the plot Yes=1, No=2	1	1.62	2	2	0.49
Animal traction	Use of animal traction on the plot. Yes=1, 2=No	1	1.63	2	2	0.48

Seed type planted	seed type planted on the plot	1010 = cowpea	1112 = shelled rice	1082 = shelled maize	-	169.63
Cropping Method	Method of cropping. 1=monocropping, 2=inter-cropping, 3=relay cropping, 4=mixed cropping, 5=alley cropping, 6=strip cropping, 7=others	1	3	4	5	1.36
Farming Age	Farming age in years	10	23	21	77	11
Sex	Male=1, Female=2	1	1	1	2	0.24
Malaria Prevalence	Malaria burden *per 1,000 per annum at 5km x 5km* grid locations	0.13	0.49	0.50	0.75	0.12
Laspeyres output	laspeyres index of total harvested crops	0	482.82	200	10000	846.25

The presence of zero value in the number of hours worked on the farm by the household might be due to the fact that the household employed hired labour to carry-out all of their farming activities. Thus, the household probably spent all of its time either on recreation or off-farm activities. At the other extreme is the household, which did not employ any hired labour in any of its farming activities. Thus, a plausible interpretation of this is that the household spent most of its time on-farm activities. Households with zero output also had zero value for total number of seeds planted.

7.2.4. How the variables are arranged for the analysis of the Nigeria data

Production Variables

Age, pesticides, equipment use, malaria prevalence, and, constant.

Inefficiency Variables

Labour, land, herbicide, animal traction, gender, and, constant.

In the next section, we discuss the Ethiopia data.

We have focussed on the Nigeria data in this subsection, in the next subsection, our focus turns to the Ethiopia data. We discuss the process of its collection and also present basic descriptive statistics on the variables.

7.2.5. The Ethiopia Data Set

The first attempt at the collection of the Ethiopia data set starts in 1989, when the International Food Policy Research Institute (IFPRI) decides to collect data from seven Peasant Associations in the regions of Amhara, Oromiya and the Southern Ethiopia People's Association (SNNPR) (Dercon and Hoddinott 2011 serves as the motivation for this section). In Ethiopia, the Peasant Association is the smallest administrative unit and it is in charge of land allocation, tenure, and distribution. A Peasant Association consists of one or more villages; as of 1990 there are over 20,000 Peasant Associations in Ethiopia (we use the word villages and Peasant Association interchangeably in this research).

The International Food Policy Research Institute end up surveying six farming villages out of the original seven intended; this is because of conflict in the seventh region of Tigray. As stated in Dercon and Hoddinott(2011 p.6), this region is inconsequential in the survey because it is mainly a semi-pastorialist region, but the survey focuses on crop farming households. The six villages are randomly selected from the chosen Peasant Associations. The Peasant Association chosen are from areas that suffered from the 1994-1985 famine and the droughts of 1987 and 1989. Four of the areas - Dinki, Koro-degaga, Gara Godo, and Doma - are considered vulnerable while the remaining two - Debre Berhan, and Adele Keke - are considered less vulnerable. The 1989 survey centres around the collection of consumption, asset, and income information of four hundred and fifty households in the six selected farming villages.

In 1994, the Centre for the Study of African Economies, University of Oxford, in conjunction with, The International Food Research Institute and the Economics Department of Addis Ababa University decide to collect longitudinal data with the 1989 data serving as a template. In other words, it will be right to say that the panel data collection actually starts in 1994. This is because the aim of the 1994 data is different from that of the 1989 data. The 1994 data set focuses on the collection of data from diverse farming households in different regions in Ethiopia cultivating different crops under different field conditions. The total number of villages increased by nine, making a total of fifteen villages surveyed.

These nine additional villages were selected with the motive of taking into account the heterogeneous nature of crop farming in Ethiopia. The consideration of the heterogeneous system of farming took precedence over administrative boundaries during stratification. The diverse systems of crop farming in Ethiopia consist of the grain-plough areas of the Northern and Central highlands, the ensette-growing areas, and the sorghum-hoe areas. Thus, the nine additional villages are from the grain-plough areas of the Northern and Central highlands, the ensette growing areas and the Sorghum-hoe areas of Ethiopia. Also, we stress that these panel survey forms strata based on the primary agro-ecological zones and sub-zones in Ethiopia and between one to three villages in are selected per strata.

As against the 1989 survey, this survey focuses on household characteristics, agriculture

and livestock information, food consumption, health, women's activities; community level data on electricity and water, sewage and toilet facilities, health services, education, Non-Governmental Organisation (NGO) activity, migration, wages, and, production and marketing. At present, seven rounds have been undertaken; two rounds in 1994, one each in 1995, 1997, 1999, 2004, and 2009. The total sample for the panel survey consists of 1,477 households from fifteen different Peasant Associations.

Selection of Sample for the Panel Survey

The Centre for the Study of African Economies and Addis Ababa University incorporated the villages used by The Internaional Food Research Institute in 1989 into 1994 survey. They used the "tracing rule" in defining panel households for the survey. They defined a household as one that still had members of the 1989 households living there - even if the head of the household had left or died. They observe that about eight percent of households had new household heads and these new heads are mostly women. They try to re-randomize the sample by replacing the households which had left the community with households with similar socio-demographic characteristics as the antecedent households. These new households could either be freshly migrated families into the village, families formed from splits or families formed from new marriages within the communities. They maintain the total number of four hundred and fifty, which was the number interviewed by The Internaional Food Research Institute in 1989. They report a low *attrition rate* (attrition rate is the number of samples lost since the last survey was conducted) of less than seven percent , which they say is most likely caused by the fact that households cannot obtain land from the Peasant Associations when they relocate to other areas.

Sampling of the whole survey was done by first obtaining a list of households in all the selected villages from the local Peasant Associations. All the households in each village were then stratified according to the sex of the household head, in other words, the households were divided into female headed and male headed households. The samples were then selected by random sampling; in all a total of 1,477 households were selected for the survey. We note that this process was also used in sampling of new households from the 1989 survey villages.

Because the different villages selected will have different population, the need for the introduction of a weighting factor is inevitable. In mitigating this effect, the team attempts linking the disputed population census figures of 1984 to the different farming systems, but they discovered that it is a very difficult process. Also, the re-delineation of administrative boundaries after the October, 1994 census is no less an important factor in the failure of this technique. In order to circumvent this, they obtain a self-weighting sample; this means that each individual selected from a village represents the same number of individuals from the primary farming systems. This method, they argue, makes the integration of the data easy. Overall, the stratification of samples enables the inclusion of representative numbers of landless households (Dercon and Hoddinott 2011 reports

that landlessness is becoming a major issue in Ethiopian agriculture). It also makes the incorporation of a fixed percentage of female headed households in the survey easy.

We reiterate that the data we use for our survey are 1994 (two rounds), 1995, 1997, 1999, 2004, and, 2009.

Merging Each of the Data Files

As stated in the survey document by Dercon and Hoddinott (2011), the 1994, 1995, 1997, 1999, 2004, and, the 2009 data are very similar in the way the questionnaires were constructed. From our point of view, the most important reason why these individual surveys were easily merged is the similar nature the data were entered in each individual file. All the files in each round are named following the *rxpxsx.dta* data entry format. They explain the meaning of this format thus:

rx refers to the round, for example, round 1 refers to the early 1994 data entered as *r1*, round 2 is the late 1994 data entered as *r2*, round 3 is the 1995 data written as *r3*, round 4 is the 1997 data entered as *r4*, round 5; the 1999 data written as *r5*, the 2004 data is the sixth round written as *r6*, and the 2009 data is written as *r7*.

px indicates the part of the questionnaire the data is for. The number of parts of the questionnaire in each round differs for some of the survey. For example, rounds 1 to 3 consist of three parts which are the 'household demographics, household assets and the non-agriculture income', 'Agriculture' and 'food consumption, health, women's activities' parts; round 4 consists of five parts which includes the first three parts of rounds 1 to 3 and two additional parts which are the 'family and marriage history, and, community and work, and, public works' parts while rounds 5, 6 and 7 consist of four parts which includes the three parts of rounds 1 to 3, and, 'shocks, public works, drought, networks, Iddir, and, trust' parts.

sx represents the section within a particular part of the questionnaire and sometimes when a particular file covers more than one section this is written, for example, as *s7t9* which means section 7 to 9.

For ease of merging our data, we construct unique household identifiers for each of the households in each round following the advise of Dercon and Hoddinott(2011, p.5) with a slight modification to their formula. This formula is written as:

$$region * 1000000 + woreda * 10000 + household\,number\,for\,the\,round\,in\,question \quad (7.1)$$

Sample Description of the Ethiopia Data set

The socioeconomic description of the Ethiopia sample reveals that the majority of the farmers in our sample have a total output of one thousand kilograms and below. This result is similar to the results of the Nigeria sample which also shows that the farmers

produced less than one thousand kilograms of output per hectare. This value is close to that of Seyoum Taffesse et al.(2011, p.8). A little above one percent of our sample have output above four thousand kilograms per hectare. This indicates that most of the farmers in our sample are small scale farmers. The Ethiopian farmers in our sample produce an average of about seven hundred and thirty kilograms per hectare (see table 7.9 and figure 7.10), this value is far below the average value we obtained for the Nigeria data. The maximum output is about ten thousand, five hundred kilogrammes per hectare, this is also very different to the results obtained from the Nigeria data.

About forty-four percent of the farmers in the Ethiopia sample are below forty years of age (see table 7.10 and figure 7.11). Thus, the farming population in our sample is young, this is a resource that needs to be harnessed, as this is not reflected in the amount of output produced by the farmers in our sample. Also, they are likely to be risk-takers and are more likely to be open to accept new innovation. About ninety-nine percent of our farmers possess one form of education or the other while about one percent do not have any form of education at all.

A lot of the farmers in our sample (about seventy percent) stop their education in their eighth grade (see table 7.11 and figure 7.12), this means that a sizeable number of our sample should be able to read and write; however, none of them have a higher education certificate. About eighty percent of the farmers in our sample operate on land size less than two hectares (see table 7.12 and figure 7.13), this means that most of the farmers in our sample are small-scale farmers. Just a little above one percent operate on land greater than five hectares. The sharing formula of the country's Peasant Association is likely to be the reason why most households operate on small-sized lands.

Fertiliser usage among the farmers is still very low this is because more than eighty-three percent of the farmers use either zero or less than one kilogram of fertiliser (see table 7.13 and figure 7.14). This is likely to have great impact on their output and the quality of farm produce. Just about seventeen percent use fertiliser above five hundred kilograms. The small amount of fertiliser used may be because most of the farmers (about eighty percent operate on good to medium fertile soils, though, in general the soil in Ethiopia is not of very high quality see Dercon and Hoddinott 2011).

The important reason why farmers source for loan is for social functions (see table 7.14 below), while, the second most important reason is to purchase items for the use of the household. About nine percent sourced for loan to buy farm inputs. Thus, a lot of the farmers in our samples are less concerned about sourcing loans for use on their farms and this may have great impact on their productivity.

Table 7.9.:   Distribution of Ethiopia Farmers by Total Output

Output (kg/ha)	Frequency	Percentage (%)
0-1000	1063	79.2
1001-2000	208	15.5
2001-3000	36	3
3001-4000	20	1.4
4001-5000	8	0.6
5001 and more	8	0.6
Total	*1343*	*100*

Figure 7.10.:   Output of Farmers in Kilograms
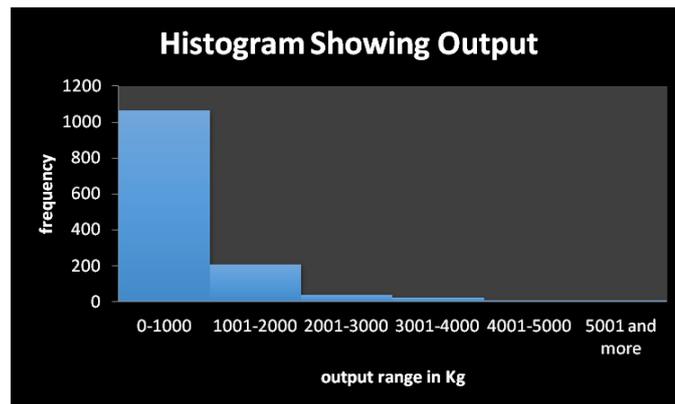
Figure 7.11.:   Age Distribution of Our Sample in Years
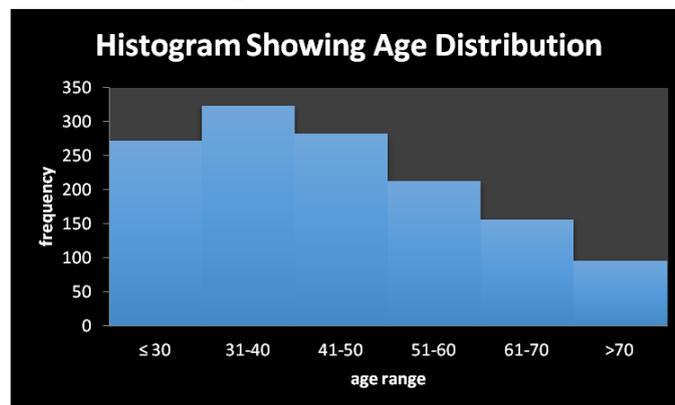
Table 7.10.:   Tabular representation of the Ethiopia Farmers Age

Age (yrs)	frequency	Percent
≤ 30	272	20
31-40	323	24
41-50	283	16
51-60	213	16
61-70	156	2
>70	96	7
Total	**1343**	**≈100**

Figure 7.12.:   Distribution of Our Sample by Education
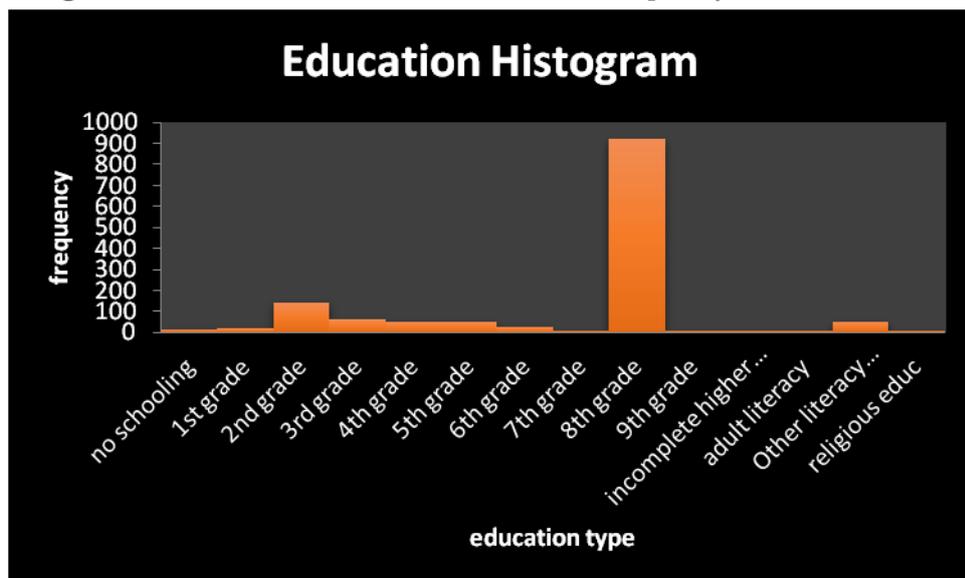
Table 7.11.:   Tabular Representation of the Farmers Education

Education	Frequency	Percent
no schooling	13	0.97
1st grade	19	1.41
2nd grade	145	10.80
3rd grade	61	4.54
4th grade	49	3.65
5th grade	48	3.57
6th grade	28	2.08
7th grade	6	0.45
8th grade	919	68.43
9th grade	1	0.07
incomplete higher educ.	1	0.07
adult literacy	3	0.22
Other literacy program	49	3.65
religious educ	1	0.07
Total	**1343**	**100**

Table 7.12.:   Table Showing total Land Cultivated in Hectares

land area (ha)	*percent*	*Frequency*
0-1	56.81	763
1.0-2.0	24.80	333
2.0-3.0	10.42	140
3.0-4.0	5.29	71
4.0-5.0	1.41	19
5.0-6.0	0.52	7
6.0-7.0	0.45	6
7.0 and above	0.30	4
Total	*100*	*1343*

Figure 7.13.:   Land Size in Hectares
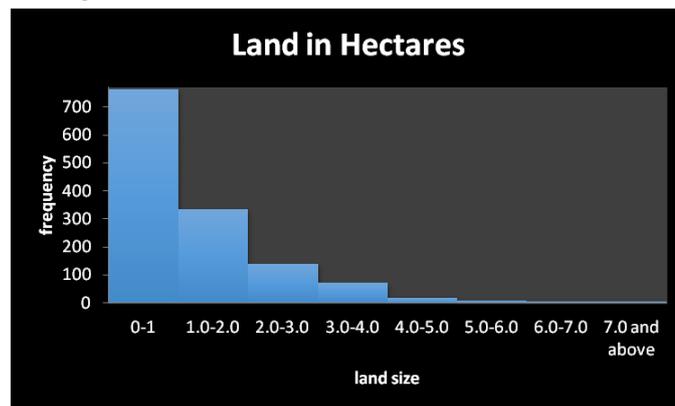
Figure 7.14.:   Total Amount of fertiliser Used in Kg/ha
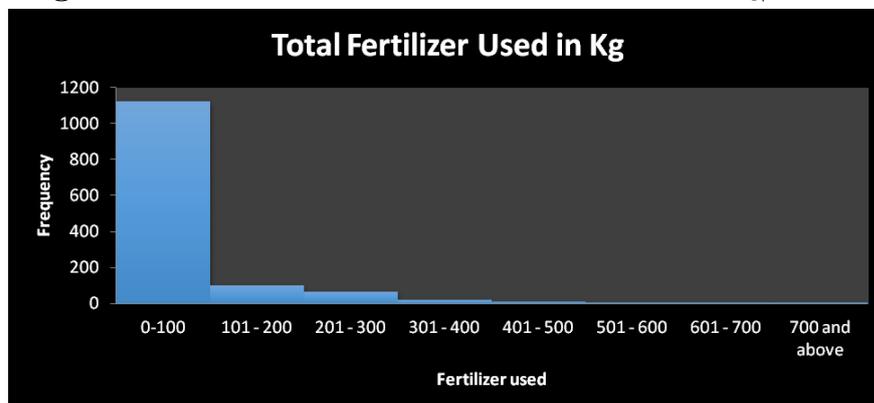
Table 7.13.:   Amount of fertiliser Used in Kg/Ha

fertiliser amount	frequency	percent
0-100	1122	83.54
101 - 200	103	7.67
201 - 300	66	4.91
301 - 400	23	1.71
401 - 500	12	0.89
501 - 600	6	0.45
601 - 700	4	0.30
700 and above	7	0.52
Total	*1343*	*100*

Figure 7.15.:   Reason for Loans
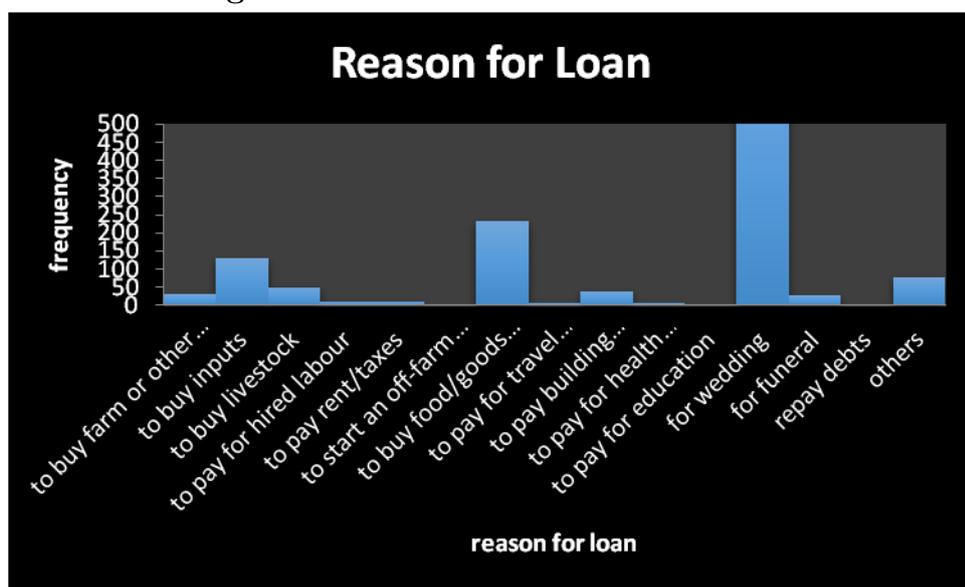
Figure 7.16.:   Quality of Land Used for Farming
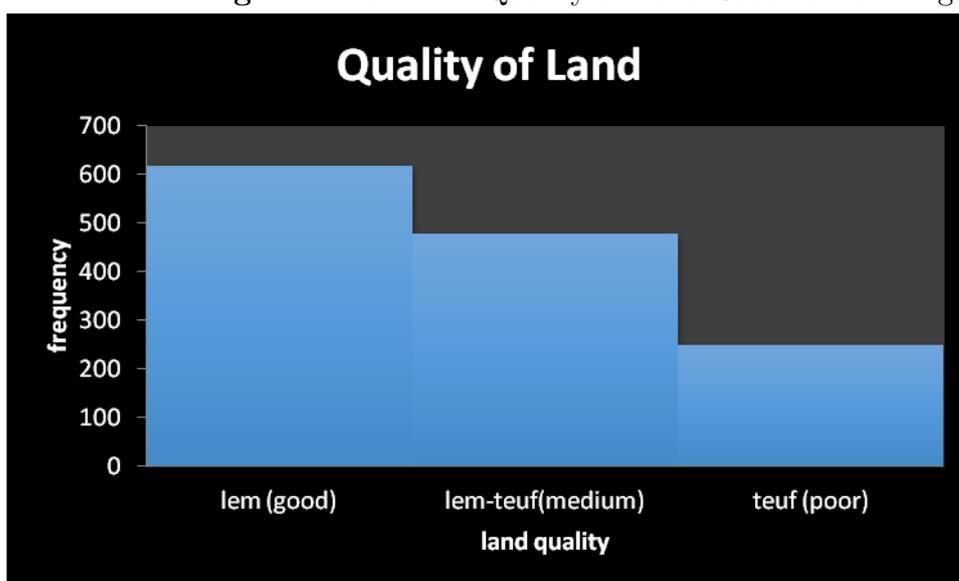
Table 7.14.:   Table Showing the Reasons for Loan

Reason for Loan	Frequency	Percent
To buy farm or other tools	32	2.38
To buy inputs	130	9.68
To buy livestock	49	3.65
To pay for hired labour	10	0.74
To pay rent/taxes	10	0.74
To start an off-farm business	0	0
To buy food/goods for household	234	17.42
To pay for travel expenses	6	0.45
To pay building material	37	2.76
To pay for health expenses	6	0.45
To pay for education	4	0.30
For wedding	715	53.24
For funeral	29	2.16
Repay debts	4	0.30
Others	77	5.73
Total	*1343*	*100*

Table 7.15.:   Quality of Land Used for Farming

Land quality	Count	Percent
lem (good)	618	46.02
lem-teuf(medium)	477	35.52
teuf (poor)	248	18.47
Total	*1343*	*100*

As usual, we present the characteristics of the variables we utilize in our analysis below:

Table 7.16.:   Characteristics of the Variables in the Ethiopia Data

S/N	Variables	Description	Min value	Mean	Median	Max	Std Dev
1	time	survey round	1	1.142	1	6	0.810
2	agg_output	laspeyres type sum of all outputs in the survey	0	41.185	1	11500	425.011
3	area_plted	total land area planted in hectares	0	1.212	0.875	9.875	1.164
4	tot_lab	total labour	0	73.003	41	10771.9	301.957
5	tot_fert	total fertlizer used	0	52.968	50	1030	112.765
6	loan_type	loan type; 99=don't know, 1=loan from money/lender/arata, 2=loan from friend/relative, 3=loan from bank, 4=other	1	5.742	8	99	4.445
7.	Land Type	Nature of soil used for farming 1=Lem (fertile), 2=Lem-Teuf (medium), 3 = Teuf (poor)	1	1.727	2	3	0.763
8	tot_input	laspeyres type index for total amount of input used	0	44.822	0	1153	95.712
9	storage	currently storing; 1=yes, 2=No	1	1.926	2	2	0.261
10	Asset	Do you own any assets 1=Yes, 2= No	1	1.020	1	2	0.15

S/N	Variables	Description	Min	Mean	Median	Max	Std.Dev
11	Reason for Loans	The reason why the farmer took loan. 1= To buy farm or other implements, 2= To buy inputs such as seeds/fertiliser, 3= To buy livestock, 4= To pay for hired labour, 5= To pay rent/taxes, 6= To start an off-farm business, 7= To buy food/goods, 8= To pay for travel, 9= To pay for building material, 10= To pay for health expenses, 11= To pay for education expenses, 12 and above= Others	1	12.593	12	112	18.447
12	credit	do you belong to any credit association, 1=Yes, 2=No	1	1.815	2	2	0.389
13	land_slope	the nature of the slope of the land, 1=medda (flat), 2=dagath-ama (gentle), 3=geddel (steep)	1	1.368	1	3	0.593
14.	Intercropp-ing	Did you intercrop 1=Yes, 2= No	1	1.719	2	2	0.453
15.	Mean Age	mean Age of the household	20	46.174	45	130	16.449
16	education	higest grade completed	0	7.047	8	19	3.220
17	mal_prev	malaria prevalence values *in case per 1000 per annum*	0.005	0.019	0.015	0.055	0.012
18.	Price	price of output per kilogram	0.25	2.946	1.670	76	4.569

7.2.6. How the variables are arranged for the analysis of the Ethiopia data

Production Variables

Asset, slope, intercropping, and, constant.

Inefficiency Variables

Land area, fertiliser, credit, malaria prevalence, and, constant

Next, we discuss the Tanzania data and its peculiarities.

7.2.7. The Tanzania Data

The LSMS-ISA survey also referred to as the National Panel Survey by Tanzania National Bureau of Statistics took off in 2009, a year earlier than that of Nigeria. The overall objective is the same as that of Nigeria, however, it has its own specific objectives which is the provision of reliable data for monitoring poverty and tracing the success of the Tanzania government's welfare scheme referred to as *Mkukuta* Poverty Reduction Strategy. The questionnaires and questions asked are similar to that of Nigeria but not exactly the same. Apart from the three different questionnaires mentioned in section 7.2.1 above, a fishery questionnaire was also included. We utilized the first round of this survey, thus, the Tanzania data we utilize in our research is cross-sectional in nature. The second round of the survey was published at the time we had finished processing the first round data and it was too late to include it in our analysis.

Data Selection and Design

As we hinted in the introductory part of this section, the data selection process is similar to that of the Nigeria data. The Tanzania National Bureau of Statistics selects the samples from a National Master Frame which, according to them, is the 2002 National Population and Housing Census. This was done in the spring of 2008. The sample for this survey also includes a sub-sample of the 2006/2007 household budget survey. In all a total of 3,280 households in 410 enumeration areas were canvassed. The breakdown of the households include 2,064 households in rural areas and 1,216 urban areas (please see Tanzania NBS , p.7 Basic Information document for more information). The Tanzania National Bureau of Statistics basic information document neither provide us much information on how the figures were arrived at nor on how the enumeration areas and households were selected.

Sample Description of the Tanzania Farmers

The socio-demographic description of the Tanzania sample shows that about sixty percent of the farmers in our sample have output index of less than four while about forty-two percent have output index of four and above (table 7.17 and figure 7.17). About sixty percent of the farmers cultivate between one and hundred hectares of land (see table 7.18 and figure 7.18). Thus, among the three countries in this research, the farmers in Tanzania cultivate on the largest piece of land. Six percent more of farmers cultivate on

land size greater than two hundred hectares than on land size between one hundred and two hundred hectares.

The level of educational attainment in our sample is very poor (see table 7.19 and figure 7.19). This is because less than one percent of the farmers in our sample finished primary school. This is expected to reflect on the level of adoption of technology, however, this has not been the case because about sixty percent of the farmers have adopted one form of erosion control technology or the other (table 7.20 and figure 7.20). About forty percent of the farmers in our sample did not adopt one form of erosion control or the other. This is unexpected because most of the farmers (about seventy eight percent) in our sample are very young. Thus, they should be more likely to adopt new innovations. Just about four percent of our sample is above fifty years of age (see table 7.21 and figure 7.21).

More than eighty percent of the farmers in our sample use land free of charge (see table 7.22 and figure 7.22 below). This is an indication of the kind of social interaction that exists among the households in our sample. About ten percent of the farmers in our sample rented the land they cultivate, while, about three percent and about 0.5 percent shared-rent and shared-owned respectively the piece of land cultivated. Intercropping is not a popular method of cropping among the households in our sample (see table 7.23 and figure 7.23). This is because less half of the households in our sample actually adopt intercropping as a cropping system. A plausible explanation for this is that most of the households in our sample grow the same class of crops that have similar soil nutrient requirements.

Figure 7.17.:   Index of Output of Crops
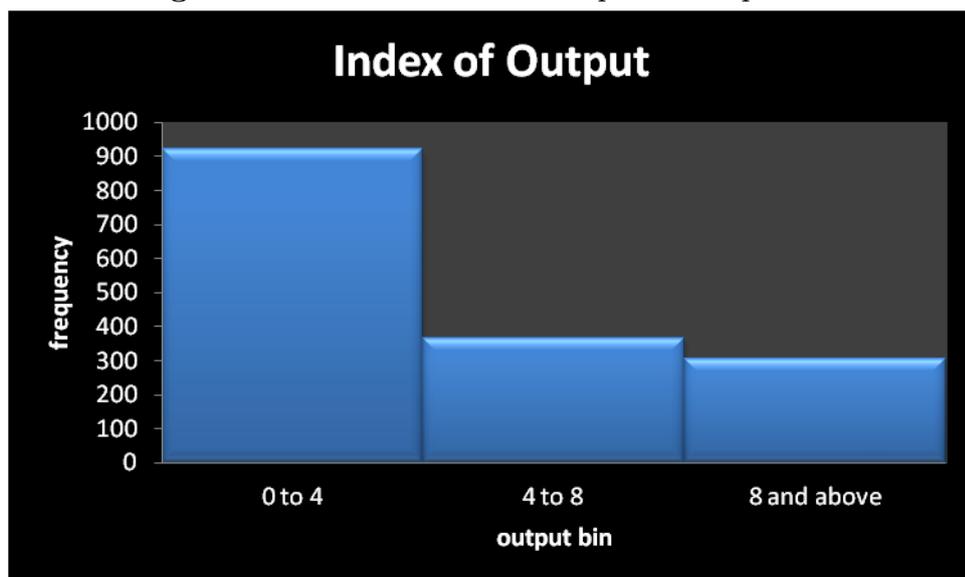
Table 7.17.:   Table Showing Index of Output

Range	Frequency	Percentage
0 - 4	923	57.9%
4 - 8	367	23%
8 and above	305	19.1%
Total	*1595*	*100%*

Figure 7.18.:   Total Land Area Grown with Crops
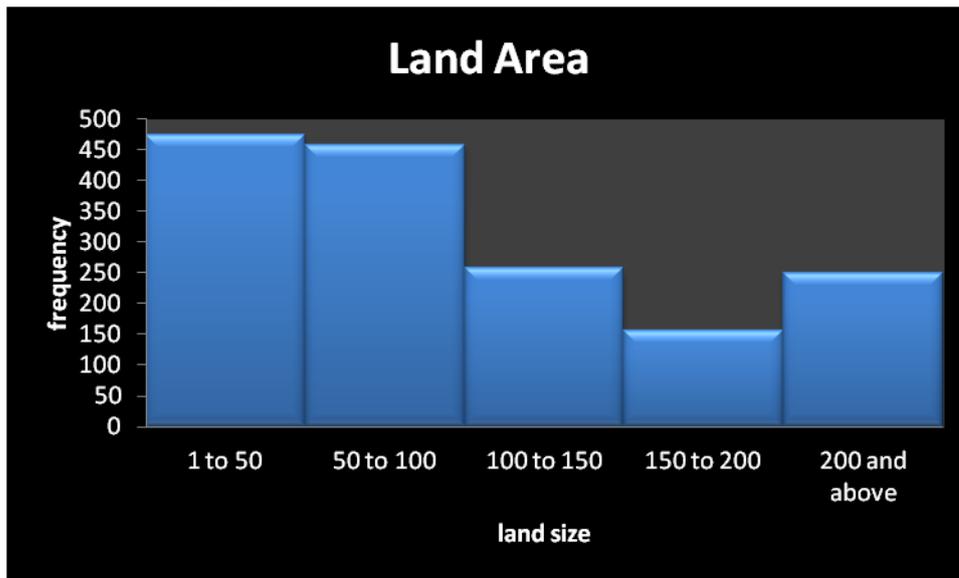
Table 7.18.:   Showing Total Land Area Cultivated in Hectares

Distribution	Frequency	Percentage (%)
1 - 50	474	29.7%
50 - 100	457	28.7%
100 - 150	258	16.2%
150 - 200	156	9.8%
200 and above	250	15.7%
Total	*1595*	*100%*

Figure 7.19.:   Showing Primary Education Completed
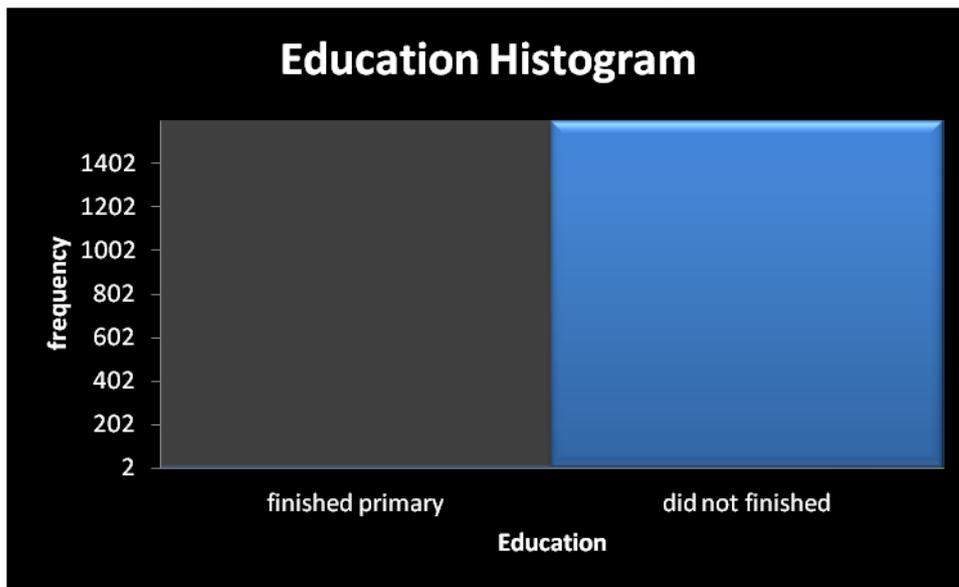
Table 7.19.:   Percentage of Farmers Who Finished Primary Education

Distribution	Frequency	Percentage (%)
finished primary	3	0.2%
did not finished	1592	99.8%
Total	*1595*	*100%*

Figure 7.20.:   Adoption of Erosion Control Technology
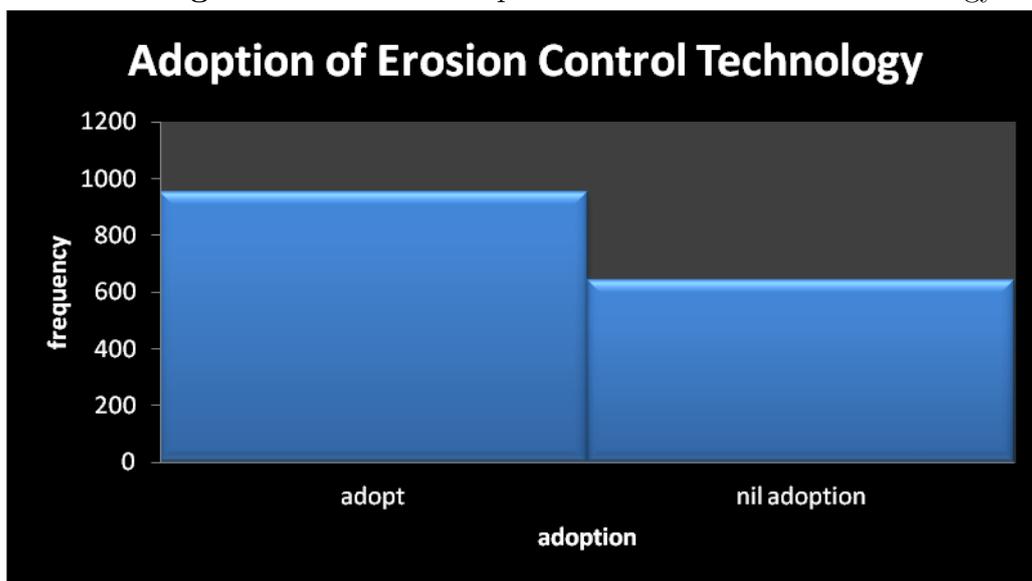
Table 7.20.:   Showing Erosion Control Technology Adoption

Distribution	Frequency	Percentage (%)
Adopt	955	60%
Nil adoption	640	40%
Total	*1595*	*100%*

Figure 7.21.:   Showing Age Distribution
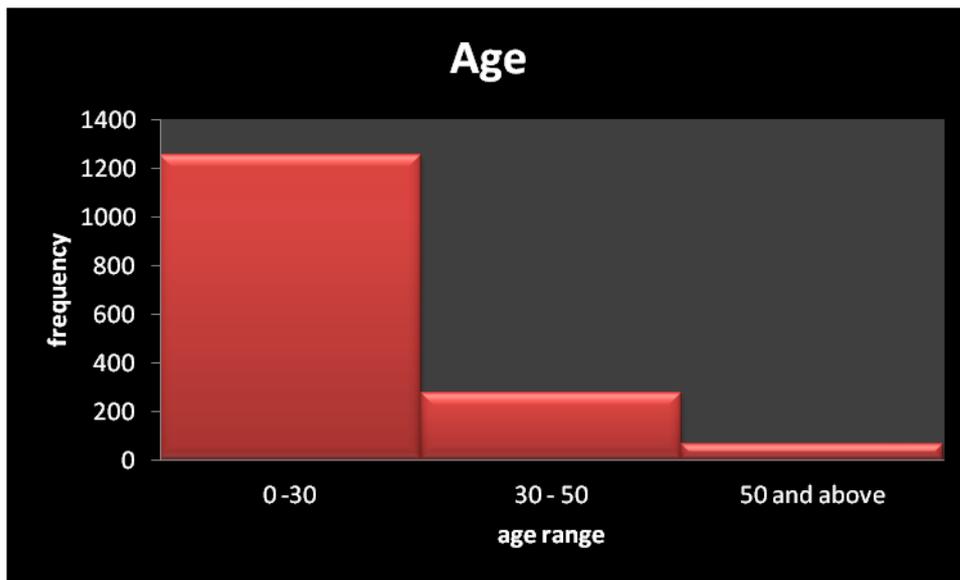
Table 7.21.:   Showing Age Distribution in Years

Distribution	Frequency	Percentage (%)
0 -30	1255	78.7
30 - 50	274	17.2
50 and above	66	4.1
Total	1595	100%

Figure 7.22.:   Land Ownership
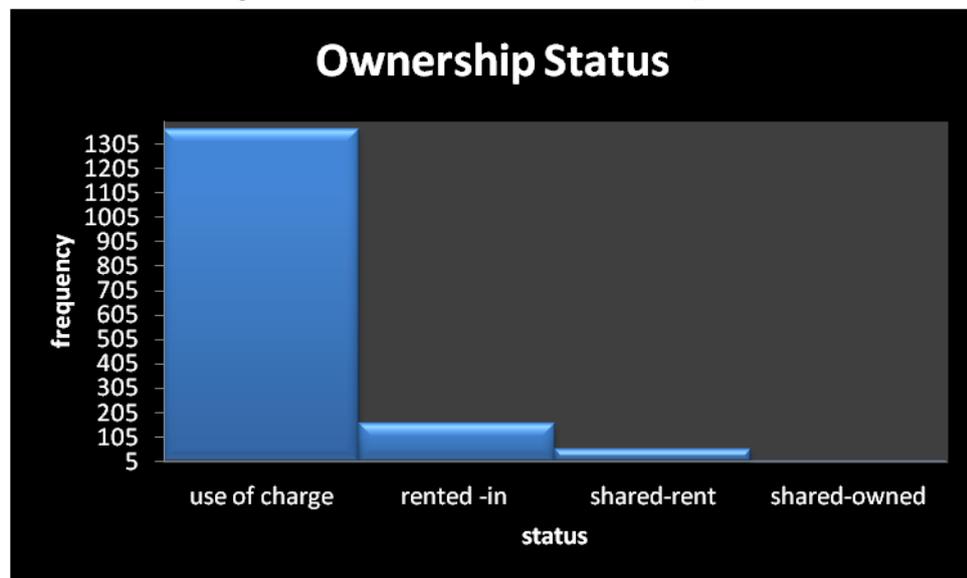
Table 7.22.:   Showing Land Ownership Distribution

Distribution	Frequency	Percentage (%)
use free of charge	1367	85.7
rented -in	162	10.2
shared-rent	56	3.5
shared-owned	10	0.63
Total	**1595**	**100%**

Figure 7.23.:   Showing Number of Farmers that Intercropped
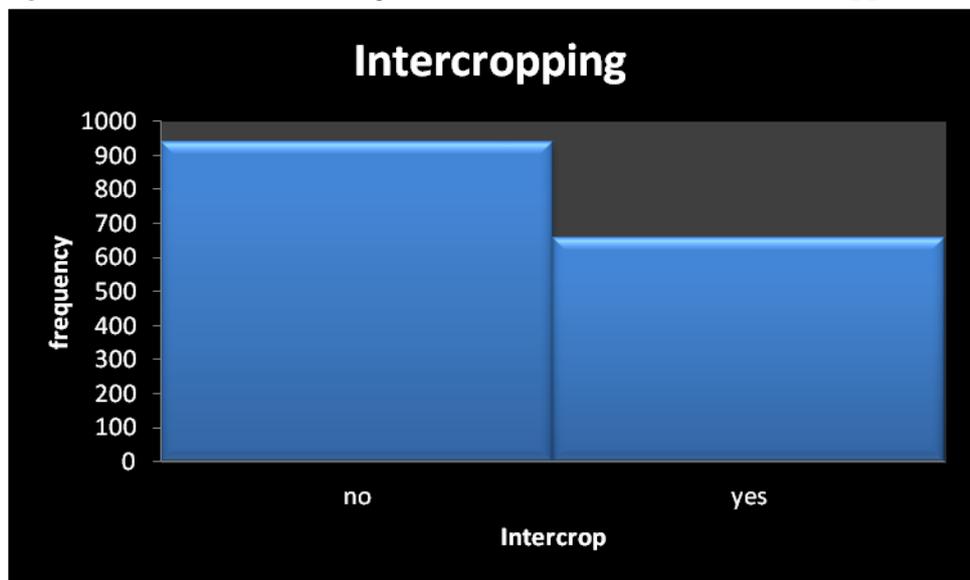
Table 7.23.:   Showing the Number of Farmers That Intercropped

Distribution	Frequency	Percentage (%)
No	938	58.8
Yes	657	41.2
Total	**1595**	**100%**

We present the characteristics of the variables we use in our analysis in the table below:

Table 7.24.:   Characteristics of the Variables in Our Sample

S/N	Variables	Description	Min	Mean	Med	Max	Std
1	harvested_qty	log of index of quantity harvested	0.69	6.22	6.29	11.28	1.35
2	area	log of index of Area cultivated	-5.30	1.21	1.25	6.44	1.05
3	labour	log of index of labour used	0.69	4.37	4.41	6.69	0.93
4	other_inputs	log of index of other inputs used	0.69	1.05	1.10	1.10	0.13
5	erosion_control	erosion control/water harvesting facility on this plot. 1=Yes, 2=No	1	1.19	1	2	0.40
6	soil_qlity	soil quality. 1= Good, 2 = Average, 3 = Bad	1	1.54	1.50	3	0.53
7	mean_educ	mean education	1	16.60	17	33	2.88
8	extension	Did you receive advice for your agricultural activities from 1=Yes, 2=No	1	1.93	2	2	0.27
9	input_credit	Seeds_fertilizers_ pesticides_herbicides received on credit to be paid later. 1=Yes, 2=No, 3=Don't Know	2	2.98	3	3	0.119
10	age	mean age	0	38.03	28.33	429	42.23
11	ownership_status	Ownership status of this plot. 1 = Owned, 2 = used free of charge, 3 = rented in, 4 = shared rent, 5 = shared ownership, 6 = others	2	2.20	2	6	0.55
12	intercropping	Was cultivation intercropped 1=Yes, 2=No	1	1.59	2	2	0.49
13	mal-prev	Malaria prevalence	0.005	0.156	0.133	0.450	0.145
14	sales	Total value of sales in Tanzanian -shillings	400	139731.09	30000	4886400	361539

P.T.O

7.2.8. How the variables are arranged for the analysis of the Tanzania data

Production Variables

Labour, erosion control, credit, age, malaria prevalence, and, constant.
Inefficiency Variables

Area planted, other inputs, extension services, malaria prevalence, constant.

This section succeeds in discussing the cross-sectional and panel data sets, in the next section, we discuss the malaria data sets and its generation.

7.3. The Malaria Data

The last section focused on the cross-sectional data sets for each of the study areas. In this section, we will explain the spatial malaria data set.

We obtained the malaria data from the Malaria Atlas Project (M.A.P) based at Oxford University. The Malaria Atlas Project (M.A.P) is principally funded by Wellcome Trust® and they collaborate with other researchers from around the world and also with international organisations like the World Health Organisation (WHO) and Measure Demographic and Health Survey (measure DHS). They produce continuous cartographic maps of malaria prevalence from around the world using Bayesian geostatistical tool to generate candidate maps at different levels of *realizations* (these are random values and also a synonym for 'candidate maps'). We decided to use the data generated from their carthographic maps as against hospital reported cases because not all infection with malaria leads to disease, especially, in areas where the population have developed immunity to the disease (Hay et al. 2010). Also, areas with high malaria prevalence have poor national malaria surveillance facilities, also malaria diagnosis is inaccurate. Their use of the Bayesian geostatistical technique to generate candidate maps enables malaria prevalence maps to be generated for areas where existing data are sparse or nonexistent like in Central Africa (see Patil et al. 2011 for more explanation). This technique also enables them to generate a predicted posterior distribution of the clinical burden of malaria. The first of these types of map was generated in 2007 and then updated in 2010, we utilize the 2010 version in our research.

Next, we highlight the method used to generate these maps.

7.3.1. Generating the Malaria Clinical Burden Cartographic Maps

The first step in obtaining acceptable cartographic maps involves the collection of initial case reports for different countries from several administrative units (Hay et al. 2010 and Gething et al. 2011 motivate this section). In total, thirteen thousand, four hundred and forty-nine administrative units were consulted. They considered eighty-five endemic countries for *Plasmodium falciparum* malaria instead of the eighty-seven considered in

2007. This is because Belize and Kyrgystan have reported zero *Plasmodium falciparum* cases since 2007.

The classification of whether a country is endemic or not is based on the classification of the Global Malaria Eradication Programme. Out of the eighty-five different countries, five (Panama, Tajikistan, Iran, Saudi Arabia, and South Africa) had dependable data for use directly to generate the cartographic maps. In the remaining eighty countries, they divided the countries into regions of stable and unstable risk of malaria transmission using the classification of the Global Malaria Eradication Programme. The Global Malaria Eradication Programme classifies a region as being a region of unstable if it reports a case of 0.1 case per 1,000 per annum. The Malaria Atlas Project data used multiplies this value by the population sizes for 2010 (the Global Rural Urban Mapping Project provides the population counts, which are then projected to 2010) and this gives the predicted malaria clinical burden for regions of unstable malaria.

To arrive at the disease burden for regions of stable malaria, they use a bespoke Bayesian geostatistical Markov Chain Monte Carlo algorithm that models the relationship between clinical incidence, prevalence and a global malaria endemicity map (this is a map that shows malaria infection risks in different countries) to generate continuous realizations which are age-standardized within the limit of the stable malaria transmission. They also measured the level of uncertainty between the extant clinical malaria incidence rate in this region and their standardised measured parasite prevalence rate. They use a Bayesian non -parametric model to achieve this objective.

The two Bayesian models - Bayesian geostatistical model and Bayesian non-parametric model - were merged together and this generates joint *realizations* of malaria attack rates for every pixel (a geographic unit which defines the resolution of a map). These malaria attack rates were then multiplied by the total pixel population to produce a predicted posterior malaria burden. They observed that uncertainty in the malaria burden in Africa showed significant differences across the different countries with high figures observed in regions with high population, but poor malaria surveillance coverage like Nigeria and the Democratic Republic of Congo. The schematic diagram in figure 7.24 below details these processes.

The summary map was generated at 5km x 5km pixels and represented as a raster (This is a grid or a layer of pixels different from vector layers which is made up of lines, points, and, polygon). The raster values were available at different percentiles, but we chose the median percentile raster. The median values of each of the three countries raster were extracted using the Geographic Information System software ESRI ArcGIS$^{©}$.

Figure 7.24.:   Flow Chart Showing How The Malaria Data Was Generated
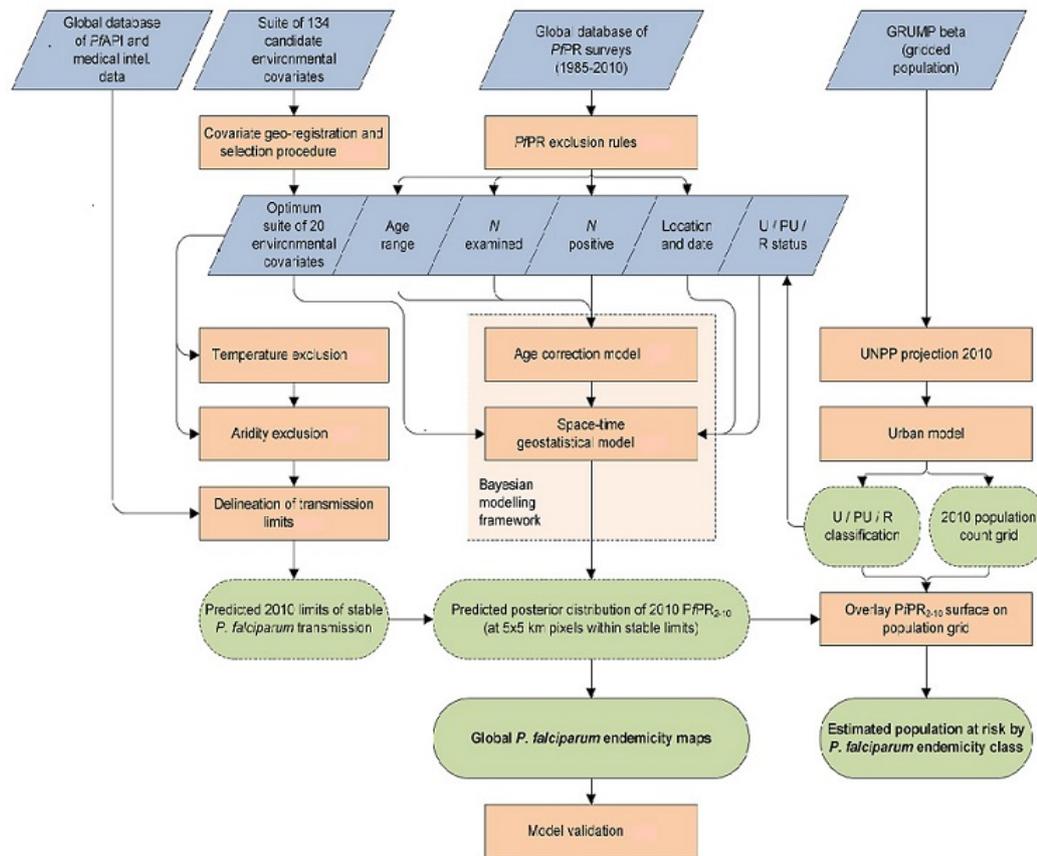
7.3.2. Extracting and Merging The Malaria prevalence Values for Each Country

The first step in extracting the median values from the Raster map was to ask for the shape file of each of our countries of interest and the corresponding Geographic Positioning System (GPS) values of each of the household in question from the custodians of the household data for each country. Afterwards, we open the raster map in ESRI ArcGIS©following these steps:

$geoprocessing \rightarrow Arctools \rightarrow conversion\,tools \rightarrow to\,raster \rightarrow float\,to\,raster$

Then, we overlay the shape file and the GPS points for each of our countries of interest on the raster maps and then extract the median malaria prevalence values following these steps:

$Data\,management\,tool \rightarrow make\,XY\,events \rightarrow spatial\,analyst\,tools \rightarrow extract\,values\,to\,points \rightarrow$
$open\,attribute\,table \rightarrow export\,as\,text\,file \rightarrow open\,text\,file\,in\,Excel^{©}$

Once the values had been exported to Excel©, this was then merged with LSMS-ISA data to obtain a combination of the LSMS-ISA data set and the malaria prevalence values.

We depict a part of the extracted raster value from the cartographic map of the Malaria Atlas Project from the ESRI ArcGIS©software in figure 7.25 below:

Figure 7.25.: A Sample of the Extracted and Merged Raster (Nigeria) Malaria Prevalence Value in ESRI ArcGIS©

Table — amgh48

s11aq5	s11aq5b	lat_dd	lon_dd	hh_lat	hh_lon	km2hh	check	hh_ch	ll_blank	ll_switc	typo	typo_	typo_lo	dups	RASTERVALU
009.01.	012.03.849	12.0641	9.01976	12.0638	10.021767	109	1	0				1			0.485
008.18.	011.25.592	11.4265	8.31256	11.257467	8.186767	23.3	1	0						1	0.485
008.18.	011.25.592	11.4265	8.31256	11.257467	8.186767	23.3	1	0						1	0.485
008.18.	011.25.682	11.4280	8.31245	11.257467	8.186767	23.4	1	0							0.485
009.12.	001.15.943	1.26571	9.20628	11.995917	9.207417	1194.1	1	0				1			0.485
009.12.	011.59.353	11.9892	9.204	11.9968	9.20665	0.9	0	0						1	0.485
009.12.	011.59.353	11.9892	9.204	11.995267	9.20775	0.8	0	0						1	0.485
004.01.	010.16.414	10.2735	4.01666	10.273567	4.683667	73	1	0				1			0.485
	010.16.165	10.2694	0	10.26945	4.68425	0	0	0	1						0.485
006.69.	005.55.139	5.91898	7.1525	5.918933	6.983467	18.7	1	0				1	1		0.485
007.47.	006.29.360	6.48933	7.78505	6.48025	7.780617	1.1	0	0						1	0.485
007.47.	006.29.360	6.48933	7.78505	6.48195	7.7816	0.9	0	0						1	0.485
007.39.	006.39.210	6.6535	7.66235	6.78375	7.662317	14.5	1	0							0.485
		0	0	6.783167	7.661833	0	0	0	1						0.485
		0	0	6.783633	7.662	0	0	0	1						0.485
007.17.	004.38.172	4.6362	7.29575	4.637467	7.301533	0.7	0	0						1	0.485
007.17.	004.38.172	4.6362	7.29575	4.6373	7.302717	0.8	0	0						1	0.485

-126.738 0.716 Decimal Degrees

So far, we have discussed the different data sets used in this research. We face different challenges in the usage of these data sets; we discuss these challenges in the next section.

7.4. Data Challenges and Limitations

We face several challenges in the use of these data sets in order to properly represent the choice statistical model in our theoretical framework. One major challenge is the presence of a lot of unanswered questions and empty spaces in them. We had the option of either using the Bayesian technique of dealing with missing observations or just deleting them. We resort to deleting these data points from our data sets because we do not have enough time to develop a code to deal with it the Bayesian way. In Ethiopia, the choosing of fifteen villages out of the numerous villages is not nationally representative, and is a major limitation of this data as we cannot make nationally representative inference from the data.

Also, there is the problem of how we represent the recursive nature (simultaneity) of the decision making process in our estimation procedure? Another challenge is the incorporation of this simultaneity into the (panel) dataset? The datasets cover several socioeconomic *realizations*, agricultural, and, community variables collected at different time periods. The different countries are different both culturally and in the time and season of planting (The secondary nature of our data is no less an important issue to consider). Consequently, we have to develop a method of properly comparing these countries. This problem is further compounded by the problem of merging the production data for individual countries with the spatial malaria geoprevalence data, which is collected by a different method and technique. We discuss how we are able to deal with some these problems in the next chapter, before then, we present the summary of this chapter.

7.5. Summary

Our data set is secondary in nature and we have presented how the household data sets and the way the samples were collected in each of the countries. We also presented how the spatial malaria prevalence data set was generated and how we merged these two different data sets together. To aid our further understanding of the data set we presented the individual countries descriptive statistics. We also raised some challenges with the use of this data set. In the next chapter, we present our empirical framework and discuss how we circumvent some of the limitations of this data set.

The next chapter focusses on the derivation of the empirical framework and the methodologies employed in our analysis.

8. Empirical Framework And Investigation

8.1. Introduction

The last chapter presents the description of the data sets used in this research. In this chapter, we present to the reader the procedures undertaken in order to estimate equation (4.13) in chapter four. This equation emphasizes the importance of the malaria prevalence data to the calculation of health inefficiency. The quantity on the right of this equation consists of the price of staples, the production technology and the efficiency variable. These last two variables are the main ingredients in the composed – error model also known as the stochastic frontier model. In other words, the composed-error model is germane to the calculation of our willingness to pay for malaria prevalence in Africa. Also, in this chapter we put into practice a lot of our remit in chapter five.

The challenges we mentioned in chapter seven are peculiarities that we have to consider and circumvent in order to arrive at 'lucid', 'robust', and 'acceptable' estimates in the composed error model. Another issue is the selection of the best and 'robust' estimates in the composed error model. We utilize the Bayesian method of composed-error calculation because of its versatility and ability to resolve some of the issues raised here and in section 7.4 above.

The chapter starts with describing the different distributions used in our model. We then proceed to present a pedagogic elucidation of our chosen model using the exponential distributions; introduce the Markov Chain Monte Carlo methods of Gibbs and Metropolis-Hastings sampling; and discuss the composed-error model comparison.

8.2. The Composed Error Model

Before we go ahead, we define the probability density functions we employ in this research. The probability density functions that we apply in this research are:

the multivariate normal probability density function (pdf),

$f^{mN}(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \equiv (2\pi)^{-m/2} \mid \boldsymbol{\Sigma} \mid^{-1/2} \exp\{-1/2(\boldsymbol{y} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$, where $-\infty < \mathbf{y} < +\infty$, $\mu \equiv (\mu_1, \mu_2, \mu_3, \ldots, \mu_m)'$ satisfies $-\infty < \mu < +\infty$ and where Σ is an m x m positive definite symmetric matrix;

the univariate truncated-normal pdf,

$$f^{TN}(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{a},\boldsymbol{b}) \equiv (2\pi)^{-1/2}\sigma^{-1/2}\exp\{-1/2\sigma^{-2}(\boldsymbol{y}-\mu)'(\boldsymbol{y}-\mu)\}/(\Phi((b-\mu)/\sigma)\Phi((a-\mu)/\sigma),$$
$$-\infty < y < +\infty; 0 < \sigma < +\infty; -\infty < a < b < +\infty;$$

the gamma pdf;

$$f^{G}(\boldsymbol{y}|\boldsymbol{\gamma},\boldsymbol{\lambda}) \equiv \Gamma(\gamma)^{-1}\lambda^{\gamma}y^{\gamma-1}\exp\{-\lambda y\}; \text{ where } \gamma, \lambda > 0$$

the inverted-gamma pdf;

$$f^{IG}(\boldsymbol{y}|\boldsymbol{\nu},\boldsymbol{s}) \equiv (2/\Gamma(\nu/2))(\nu s^2/2)^{\nu/2}(1/\sigma^{\nu+1})\exp\{-\nu s^2/2\sigma^2\};$$

the exponential pdf (which is the joint distribution assumed for our inefficiency)

$$f^{E}(\boldsymbol{y}|\boldsymbol{\lambda}) \equiv \lambda\exp(-\lambda y).$$ The exponential pdf and the inverted gamma pdf are variants of the gamma pdf,

the uniform pdf,

$$f^{u}(\boldsymbol{y}|\boldsymbol{a},\boldsymbol{b}) \equiv (\boldsymbol{b}-\boldsymbol{a})^{-1}, a \leq y \leq b;$$

the multivariate-t distribution,

$$f^{mT}(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{V},\boldsymbol{v},\boldsymbol{m}) \equiv \frac{v^{v/2}\Gamma[(v+m)/2]|V|^{1/2}}{\pi^{m/2}\Gamma(v/2)}\left[v+(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{V}(\boldsymbol{y}-\boldsymbol{\mu})\right]^{-(m+v)/2} \quad -\infty < y_i <$$
$\infty, i = 1,2,\ldots,m$ where $v > 0$, V is an $m \times m$ positive definite matrix, and, $\mu' = (\mu_1, \mu_2, \ldots, \mu_m)$ with $-\infty < \mu_i < \infty$, $i = 1, 2, \ldots, m$;

We will often use the symbol 'α' to indicate the proportionality of the variable part of the target density to a true pdf.

Thus, we write the truncated-normal pdf as $f^{TN}(\boldsymbol{y}|\boldsymbol{\mu},\sigma,a,b)\,\alpha\,\exp\{-1/2\sigma^{-1/2}(\boldsymbol{y}-\mu)'(\boldsymbol{y}-\mu)\}/(\Phi((b-\mu)/\sigma)\Phi(((a-\mu)/\sigma).$

In other words, the truncated Normal is proportional to '$\exp\{-1/2\sigma^{-1/2}(\boldsymbol{y}-\mu)'(\boldsymbol{y}-\mu)\}/(\Phi((b-\mu)/\sigma)\Phi(((a-\mu)/\sigma)$'.

In section 8.2.1, we describe the stochastic frontier model using the cross-section data, this is because two of our data (the Nigeria and Tanzania data) are cross-sectional in nature. Afterwards, in section 8.3 we introduce the panel data form of our model because the Ethiopia data is panel in nature, we have explained the advantage of using the panel data over the cross-sectional data also in section 5.3 above

We also assume a log-linear production function for our analysis (this assumption is based on the researcher and the prevailing economic theory, though, the non-imposition of this assumption does not preclude Bayesian estimation). The assumptions made in this chapter are usual with studies like ours and have far-reaching foundation in the literature which we have already perused in chapters three and five. In the next section, we present the Bayesian stochastic frontier model using the exponential distribution.

8.2.1. Bayesian Stochastic Frontier Model Assuming the Exponential Distribution For Inefficiency

Our aim in this section is to introduce our model for the analysis starting with a basic cross-sectional model and later extend it to the panel data model (Holloway et al. 2005 motivate a lot of our remit in this session). We present a pedagogic representation of our model using the proper prior in our analyses. This approach enables the budding Bayesian econometrician to appreciate the concept of Bayesian econometrics more as it shows a step by step process by which we arrive at our result (see Holloway, 2005 for non-informative prior presentation of this).

Our model assumes a normal distribution for the random error and an exponential joint distribution for the inefficiency term, thus, we use the normal-exponential model of choice. The exponential model places a constraint on the shape of the inefficiency distribution since it is a single parameter distribution. One advantage of the exponential distribution is its ease of tractability.

We write a generic equation for the production frontier facing the household as (please see chapter seven for a full description of all the variables we used for each country):

$$\boldsymbol{Y}_i = f(\boldsymbol{X}_i; \boldsymbol{\beta})\boldsymbol{\theta}_i^{'} \tag{8.1}$$

Where Y_i denotes the output produced by individual household unit, i; $f(\textbf{.})$ is the production function, X_i are the variables conjectured to influence the production function for household unit i, β is the parameter associated with each variable, and, $\boldsymbol{\theta}_i^{'}$ is a measure of the inefficiency assocaied with household unit i.

If we include the normally distributed measurement error as used in Koop (2003):

$$\boldsymbol{Y}_i = f(\boldsymbol{X}_i; \boldsymbol{\beta})\boldsymbol{\theta}_i^{'}\boldsymbol{\zeta}_i \tag{8.2}$$

where $\zeta_{\textbf{i}}$ is the normally distributed measurement error

Log-linearizing production equation (8.2) gives:

$$\boldsymbol{lnY} = \boldsymbol{f(x_i; \beta)} + \boldsymbol{ln(\zeta_i)} + \boldsymbol{ln(\theta_i^{'})} \tag{8.3}$$

Further simplification of equation (8.3) gives:

$$y_i = x_i'\beta + u_i - z_i \tag{8.4}$$

where $y_i = lnY$ denotes a logged $n \times 1$ vector of output; $x_i' \equiv (x_{11}, x_{22}, \ldots, x_{nk})'$ is a logged $n \times k$ matrix of factors that affect the production frontier, $\beta \equiv (\beta_1, \beta_2, \ldots, \beta_k)$ is a $k \times 1$ vector of the magnitude of the relationship between the frontier and the factors, x_i', $u_i = ln(\zeta_i)$ denotes the $n \times 1$ vector of the two-sided (stochastic) random error and $z_i = -ln(\theta_i')$ is a $n \times 1$ vector that denotes the deviation of actual output of household from the maximum feasible output (the frontier). All of these quantities have been explicitly defined for each country in chapter seven and we implore the reader to refer to them for further understanding of our model.

As stated earlier, we assume that the sampling error is normally distributed (N); the inefficiency term z_i is positive. Higher values of z_i denote greater deviation from the frontier while lower values of z_i denotes reduced deviation from the frontier. If we assume an exponential distribution (E) for the inefficiency parameter, the data density becomes:

$$f(y_i|\theta) = f^N(y_i|x_i'\beta - z_i) \cdot f^E(z_i|\lambda) \tag{8.5}$$

Equation (8.5) shows the data density as the product of the conditional normal distribution (N) and the exponential (E) distribution for the inefficiency variable. We observe y_i (the output) and x_i (the variables); but β, σ (the scale factor for the stochastic error term, u_i). We emphasize here that the scale parameter σ is same as the scale parameter represented as variance, σ^2, which is used in some econometric textbooks except that in our case we have used the square-root counterpart of the variance, σ^2, called the standard error, σ, in statistics.

The quantities z_i, and, λ, the mean value of z_i are all unobservable . We digress a bit to explain that the mean of the inefficiency parameter, λ, is effectively defined following Battese and Coelli (1995, pp. 326 - 327) such that

$$\lambda = W_i\delta \tag{8.6}$$

where: $W_i \equiv (W_1, W_2, \ldots, W_k)$ denotes covariates that affect conditional mean inefficiency with $i = 1, 2, ..., k$ and not a normalising constant (A normalising constant has a probability density function that integrates to one). If it is a normalising constant then we do not need to derive its distributional form; $\delta \equiv (\delta_1, \delta_2, \ldots, \delta_k)$ denotes how much these covariates influence technical inefficiency.

We treat the regression parameters (β and σ), z_i, and δ as unknowns, the question is; on which one of these should we place our prior? We will place our prior on β, σ, and δ, this is because we prefer to view z_i as 'latent' or 'missing' data. Hence, we treat z_i as if it is not part of the model parameter and so its prior and posterior do not enter into the calculation as well as into the marginal likelihood calculation (Holloway et al. 2005 shed more light on this). As a result, we write our model parameters as $\boldsymbol{\theta} \equiv (\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\delta})'$ and then $\boldsymbol{z}_i \equiv (z_1, z_2, \ldots, n)'$ where $i = 1, 2, \ldots, n$. As noted in Koop (2003), $z_1, z_2, \ldots, z_n$ are *a priori* independent likewise $\boldsymbol{y}_i \equiv (y_1, y_2, \ldots, y_n)$. The complete data density is:

$$f(\boldsymbol{y_i}|\boldsymbol{\theta}, \, \boldsymbol{z_i}) = \prod_{i=1}^{n} f(\boldsymbol{y_i}|\boldsymbol{\theta}, \boldsymbol{z}) \tag{8.7}$$

As stated earlier, $\boldsymbol{y}_i$ is observable while θ is not, so we place a prior density on θ, $\pi(\theta)$; and our posterior becomes:

$$\pi(\boldsymbol{\theta}|\boldsymbol{y_i}) \, \alpha \, f(\boldsymbol{y_i}|\boldsymbol{\theta}, \, \boldsymbol{z}) \centerdot \pi(\boldsymbol{\theta}) \tag{8.8}$$

Equation (8.8) states the posterior density as being the product of the data density (likelihood function) and the prior density. In our analysis, we utilize a proper prior which is of the same form as the data density. In other words, we utilize a normal pdf prior for our analysis. The literature refers to this type of prior as conjugate prior. As a result, equation (8.8) gives (we utilize 1 and 2 to distinguish between the two samples; 1 is the prior and 2 is the likelihood function)

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \, \alpha \, \prod_{i=1}^{n} [f^N(\boldsymbol{y_1}|\boldsymbol{x_1'}\boldsymbol{\beta} - \boldsymbol{z_1}, \boldsymbol{\sigma^2 I_n}) \cdot \boldsymbol{f}^N(\boldsymbol{y_2}|\boldsymbol{x_2}'\boldsymbol{\beta} - \boldsymbol{z_2}, \boldsymbol{\sigma^2 I_n})] \cdot \boldsymbol{f}^E(\boldsymbol{z}|\boldsymbol{W\delta}) \tag{8.9}$$

The quantity I_n is an $n \times n$ unit matrix

One of the several exercises econometricians carry-out is to find the posterior marginal distribution (this is the pdf of the parameter conditioned on the data **y**) of each of the parameters (see Zellner 1971 p.67 for example). This is done by integrating out one or more of the parameters in equation (8.9). The posterior marginal distributions of the model parameters are written as $\pi(\boldsymbol{\beta}|\boldsymbol{y})$, $\pi(\boldsymbol{\sigma}|\boldsymbol{y})$, $\pi(\boldsymbol{\delta}|\boldsymbol{y})$. However, one or more of the marginal distribution are not readily available and the econometrician says they are not in their 'closed' form. This means it is very difficult to integrate out one or more of the choice parameters either because the process is tedious, uneconomical, and/or it is almost impossible to do so.

At present the task in our hands is to find the posterior marginal distribution of our model parameters. If we recall equation (8.5) and exclude the variable '–z' in the equation, we have an equation looking like the normal linear regression model with model parameters, β and σ. In the normal linear regression model, the posterior marginal distribution of $\pi(\boldsymbol{\beta}|\boldsymbol{y})$ and $\pi(\boldsymbol{\sigma}|\boldsymbol{y})$ are in their closed form and they are both of the multivariate-t and inverted-Gamma distributions respectively.

Similarly, the marginal distribution of $\pi(\boldsymbol{\beta}|\boldsymbol{y})$ and $\pi(\boldsymbol{\sigma}|\boldsymbol{y})$ in equation (8.9) are in their closed form and they are of the Multivariate-t distribution and the inverted-Gamma distribution. However, it is difficult to obtain the closed form for $\pi(\delta|y)$ because its conditional distribution $\pi(\boldsymbol{\delta}|\mathbf{y}, \mathbf{z})$ is dependent on z and z is also dependent on y. Integrating out z is difficult which is the 'dilemma' we face in order to be able to obtain robust estimates for δ. The Markov chain Monte Carlo (MCMC) method(s) of Gibbs sampling (and the Metropolis – Hasting) is employed in resolving this problem. We focus on this method in the next section.

8.2.2. Markov Chain Monte Carlo (MCMC or MC^2) Methods

Posterior simulation is widely used in Bayesian econometrics. Methods of posterior simulation include rejection sampling, Monte Carlo integration, importance sampling, Gibbs sampling and the Metropolis-Hastings method. Posterior simulation involves the judicious use of the weak laws of large numbers and the (pseudo)random number generators from specified probability distributions. These methods are called Monte Carlo methods in statistics named after a popular casino in France. However, only the latter two – Gibbs sampling and Metropolis-Hastings – qualify to be called Markov chain Monte Carlo methods . This is because a 'draw' from a (marginal) distribution is dependent on preceding 'draws'. Thus, both of them would always need the researcher to specify the starting values which may later be discarded called "burn-ins". Our choice of the MCMC methods in this research is based on evidence provided in the literature and the fact that they provide less tedious (as will be observed soon) and efficient simulations.

In the rest of this chapter, we focus our attention on these two methods and for further readings on the Monte Carlo methods we recommend Gentle (2003), Gilks et al. (1996), Robert and Casella (2004), Koop (2003), Del Moral et al. (2006), Gelman et al. (2003), and, Geweke (1999).

Monte Carlo Method of Gibbs Sampling the Composed Error Model Using the Exponential Inefficiency Distribution

The process by which we arrive at the forms of the different conditional distributions is always a source of concern for budding econometricians; this necessitates our description of the procedure in this section. We employ the natural conjugate prior for this analysis.

At this juncture, we would like to remind the reader of the two conditions before one can Gibbs sample. The necessary condition is that the conditional distributions of the

model parameters in terms of the joint posterior must be in the closed form, while the sufficient condition is that the form of the conditional distributions must be tractable, that is, we must be able to draw samples from them. In the following sections, we describe the true forms of the model parameters, σ, β, z, and, δ respectively and then compare our derivations to standard distribution in the literature, we use Zellner (1971 pp. 363 – 378) as our reference in this exercise. We recall equation (8.9) and write it in the form (ignoring the capital pi sign):

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{z})\, \alpha\, [f^N(\boldsymbol{y_1}|\boldsymbol{x_1'}\boldsymbol{\beta} - \boldsymbol{z_1},\, \boldsymbol{\sigma^2 I_n})\cdot f^N(\boldsymbol{y_2} \mid \boldsymbol{x_2'}\boldsymbol{\beta} - \boldsymbol{z_2},\, \boldsymbol{\sigma^2 I_n})]\cdot f^E(\boldsymbol{z}|\boldsymbol{W\delta}) \quad (8.10)$$

Apart from $\boldsymbol{I}_n$ which is the n-dimensional identity matrix previous definitions of our variables remain the same. The sequence in which this process is carried-out is of utmost importance because a draw depends on the previous draw until the process comes to completion. The general principle of arriving at the conditional distribution of our model parameters is to first integrate out our choice parameter from the joint distribution and then compare the result to standard distributions in the literature, we can then use the Gibbs sampler or the Metropolis-Hastings algorithm to draw our samples.

The Conditional Distributions For The Parameters of Choice

The first conditional distribution we derive is that of the scale parameter, σ for the stochastic error term, u_i. A careful look at the posterior distribution, equation (8.11), shows that the scale parameter (or standard error), σ, is only present in the first part of the right hand part of the distribution, which is the multivariate-normally (N) distributed part of the distribution. Consequently, the exponential (E) part of equation (8.10) drops-out and we are left with the multivariate-normally distributed part. We also leave out of the distribution any term that does not contain σ. We are left with the expression:

$$\pi\left(\boldsymbol{\sigma} \mid \boldsymbol{\beta},\, \boldsymbol{z},\, \boldsymbol{y}\right) \alpha\, f^N(\boldsymbol{y_1} \mid \boldsymbol{x_1'}\boldsymbol{\beta} - \boldsymbol{z_1},\, \boldsymbol{\sigma^2 I_n})\cdot f^N(\boldsymbol{y_2} \mid \boldsymbol{x_2'}\boldsymbol{\beta} - \boldsymbol{z_2},\, \boldsymbol{\sigma^2 I_n}) \qquad (8.11)$$

when written in more explicit form, it becomes:

$$\alpha\sigma^{-(n_1+n_2+1)}\cdot\exp\left\{-1/2\sigma^{-2}[(\boldsymbol{y_1} - \boldsymbol{x_1}'\boldsymbol{\beta} + \boldsymbol{z_1})'(\boldsymbol{y_1} - \boldsymbol{x_1}'\boldsymbol{\beta} + \boldsymbol{z_1})] + [(\boldsymbol{y_2} - \boldsymbol{x_2}'\boldsymbol{\beta} + \boldsymbol{z_2})'(\boldsymbol{y_2} - \boldsymbol{x_2}'\boldsymbol{\beta} + \boldsymbol{z_2})]\right\}$$

$$(8.12)$$

further simplification results in:

$$\pi(\boldsymbol{\sigma}|\boldsymbol{\beta}, \boldsymbol{z},\, \boldsymbol{y})\, \alpha\, \boldsymbol{\sigma}^{-(n_1+n_2+1)} \exp\{-1/2\sigma^{-2}[(\boldsymbol{y_2} - \boldsymbol{x_2}'\boldsymbol{\beta} + \boldsymbol{z_2})'(\boldsymbol{y_2} - \boldsymbol{x_2}'\boldsymbol{\beta} + \boldsymbol{z_2})] + \boldsymbol{v_1 s_1^2}\} \quad (8.13)$$

A like for like Comparison of equation (8.14) with the inverted-Gamma distribution in Zellner, (1971 p. 371 equation A.37b) shows that $v + 1 = n_1 + n_2 + 1$ (the number of degrees of freedom) and $vs^2 == \boldsymbol{v_1 s_1^2} + (\boldsymbol{y_2} - \boldsymbol{x_2}'\boldsymbol{\beta} + \boldsymbol{z_2})'(\boldsymbol{y_2} - \boldsymbol{x_2'}\boldsymbol{\beta} + \boldsymbol{z_2})$. Hence, the prior distribution we employ for the standard error, σ is of the inverted-Gamma distribution, written for short as $\pi(\boldsymbol{\sigma}|\boldsymbol{\beta}, \boldsymbol{z}, \boldsymbol{y}) \equiv f^{IG}(\boldsymbol{\sigma}|\boldsymbol{\nu_1}, \boldsymbol{s_1^2})$, defined as $v_1 = 1,\ s_1^2 = 1$. The quantities v and s are the mean and variance of inverted-Gamma distribution (we define all β quantities in the next section). This means that the draw for σ will be from the inverted-Gamma distribution. There are several algorithms for drawing from the Inverted-Gamma distribution; one of such algorithms is to draw from a scaled *Chi-Square* distribution (See Gelman et al. 2003 p. 580).

Our next task is to find the conditional distribution form of the prevailing technology parameter, β. We integrate out β from equation (8.11) and make it the subject of the equation. Because β is only found in the regression part of equation (8.10), once again, we leave out the exponential part of the equation and other terms that has nothing to do with β (this includes preceding constants in the kernel). Hence, we have an equation of the form:

$$\pi\left(\boldsymbol{\beta} \mid \boldsymbol{\sigma},\ \boldsymbol{z}\ \boldsymbol{y}\right) \alpha\ f^N\left(\boldsymbol{y_1} \mid \boldsymbol{x_1'}\boldsymbol{\beta} - \boldsymbol{z_1},\ \boldsymbol{\sigma^2 I_n}\right) \cdot f^N(\boldsymbol{y_2} \mid \boldsymbol{x_2'}\boldsymbol{\beta} - \boldsymbol{z_2},\ \boldsymbol{\sigma^2 I_n}) \tag{8.14}$$

$$\pi\left(\boldsymbol{\beta} \mid \boldsymbol{\sigma},\ \boldsymbol{z}\ \boldsymbol{y}\right) \alpha\ \exp\left\{-\text{\small{1/2}}\sigma^{-2}[(\boldsymbol{y_1} - \boldsymbol{x_1}'\boldsymbol{\beta} + \boldsymbol{z_1})'(\boldsymbol{y_1} - \boldsymbol{x_1}'\boldsymbol{\beta} + \boldsymbol{z_1})] + [(\boldsymbol{y_2} - \boldsymbol{x_2}'\boldsymbol{\beta} + \boldsymbol{z_2})'(\boldsymbol{y_2} - \boldsymbol{x_2}'\boldsymbol{\beta} + \boldsymbol{z_2})]\right\} \tag{8.15}$$

we let $a_1 = \boldsymbol{y_1} + \boldsymbol{z_1}$ and $\boldsymbol{a_2} = \boldsymbol{y_2} + \boldsymbol{z_2}$, we write equation (8.15) as:

$$\alpha\ \exp\{-\text{\small{1/2}}\sigma^{-2}[(\boldsymbol{a_1} - \boldsymbol{x_1}'\boldsymbol{\beta})'(\boldsymbol{a_1} - \boldsymbol{x_1'}\boldsymbol{\beta}) + (\boldsymbol{a_2} - \boldsymbol{x_2}'\boldsymbol{\beta})'(\boldsymbol{a_2} - \boldsymbol{x_2'}\boldsymbol{\beta})]\} \tag{8.16}$$

On completing the square in β gives:

$\alpha\{\beta' C_{\hat{\beta}}\beta - 2\beta'(x_1'a_1 + x_2'a_2)\}\ \alpha\{-(a_1'x_1 + a_2'x_2)C_{\hat{\beta}}^{-1}(x_1'y_1 + x_2'y_2) + [\beta - C_{\hat{\beta}}^{-1}(x_1'y_1 + x_2'y_2)]'[\beta - C_{\hat{\beta}}^{-1}(x_1y_1 + x_2y_2)]\}$

we have:

$$\alpha\ \exp\{-\text{\small{1/2}}\sigma^{-2}[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'C_{\hat{\beta}}^{-1}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})]\} \tag{8.17}$$

where $\hat{\beta} = C_{\hat{\beta}}(\sigma^{-2}x_2'a_2 + C_{\hat{\beta}_1}^{-1}\hat{\beta}_1) = C_{\hat{\beta}}(\sigma^{-2}x_2'a_2 + x_1a_1)$ and $C_{\hat{\beta}} = (\sigma^{-2}x_2'x_2 + C_{\hat{\beta}_1}^{-1})$;
$C_{\hat{\beta}_1} = \sigma^2(x_1'x_1)$ and $\hat{\beta}_1 = (x_1'x_1)^{-1}x_1a_1$

Thus:

$$\pi\left(\beta \mid \sigma, z\, y\right) \alpha \, \exp\{-\tfrac{1}{2}[(\beta - \hat{\beta})'C_{\hat{\beta}}^{-1}(\beta - \hat{\beta})]\} \tag{8.18}$$

This form is same as that in Zellner (1971 p. 379 equation B.1), hence, β is of the multivariate-normal form written as $\pi(\beta \mid \sigma, z\, y) \equiv f^{MN}(\beta|\hat{\beta}_1, C_{\hat{\beta}_1})$. We utilize $\hat{\beta}_1 = 0_k$ and $C_{\hat{\beta}_1} = I_k * 1000$

Next, we find the form of the inefficiency variable, z.

The parameter, z , is found on all sides of the distributions in equation (8.10). Thus, finding its conditional distribution is more difficult than the previous two parameters. We reiterate that we did not place any prior on z and thus we will eliminate the subscript *1* and *2* and focus on just the likelihood function.

$$\pi(z|\sigma, \beta, y, \lambda) \alpha \, f^N(y|x'\beta - z, \sigma^2 I_n) \cdot f^E(z|\lambda) \tag{8.19}$$

A more elaborate presentation of equation (8.19) with respect to the parameter z is given as:

$$\pi(z|\sigma, \beta, y, \lambda) \alpha \, \exp\{-\tfrac{1}{2}\sigma^{-2}(y - x'\beta + z)'(y - x'\beta + z)\} \cdot \{\exp(-\lambda z)\} \tag{8.20}$$

In order to arrive at the appropriate conditional distribution form of equation (8.20), we 'complete the square' in z as we did in the case of β. To ease the process of completing the square in z, we re-write equation (8.21) as:

$$\pi(z|\sigma, \beta, y, \lambda) \alpha \, \exp\{-\tfrac{1}{2}\sigma^{-2}(A + z)'(A + z)\} \cdot \{\exp(-\lambda z)\} \tag{8.21}$$

where $A \equiv y - x'\beta$

Expanding the right hand side of equation (8.21) gives:

$$\alpha \, \exp-\tfrac{1}{2}\sigma^{-2}(z'z + z'A + A'z) - \tfrac{1}{2}\lambda'z - \tfrac{1}{2}z'\lambda$$

Further simplification gives:

$$\alpha \exp\{-1/2\sigma^{-2}(z'z + z'A + A'z + \sigma^2\lambda'z + \sigma^2 z'\lambda)\}$$

$$\alpha \exp\{-1/2\sigma^{-2}(z'z + z'(A + \sigma^2\lambda) + (A' + \sigma^2\lambda')z)\}$$

$$\pi(\boldsymbol{z}|\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{y}, \boldsymbol{\lambda}) \, \alpha \exp\{-1/2(z - \hat{z})C_{\hat{z}}^{-1}(z - \hat{z})\} \tag{8.22}$$

This mathematical operation gives us a distribution for z of the form, $\pi(z|\sigma, \beta, \boldsymbol{y}, \lambda) \, \alpha \, f^{TN}(\boldsymbol{z}|\hat{\boldsymbol{z}}, \boldsymbol{C}_{\hat{z}})$, which is the truncated normal distribution to $z > 0$ with parameters defined as $\hat{\boldsymbol{z}} = \boldsymbol{C}_{\hat{z}}(-\boldsymbol{A} - \boldsymbol{\sigma^2\lambda}) = \boldsymbol{C}_{\hat{z}}(\boldsymbol{x\beta} - \boldsymbol{y} - \boldsymbol{\sigma^2\lambda}), \boldsymbol{C}_{\hat{z}} = \boldsymbol{\sigma^2 I_N}$

One of the ways of drawing from the truncated-Normal distribution is to first draw from the normal distribution and then discard draws which are less than zero, that is, $z < 0$. Alternatively, one can use the Chib (1992) method; which is more efficient since it produces sample draws from the distribution without discarding any at each run (referred to as 'one-for-one' draws in Holloway et al. 2005); we make use of this approach.

The approach makes use of the *probability integral transformation* of the uniform distribution (see Mood et al. 1974 p. 202 for further explanation). This involves finding the inverse of the cumulative distribution of a uniform distribution over the interval (a, b). To do this, we obtain a scalar value say τ from the standard normal distribution over the interval 'a' and 'b', such that $\tau = \Phi^{-1}((U(a,b))(\Phi\left(\frac{\mu(b)}{\sigma}, 1\right)))$, where $\Phi(.)$ is the standard Normal cumulative distribution function, Φ^{-1} is the inverse of the standard normal cumulative distribution function, μ is the mean which is given by $x'\beta$. Then z equals a multiplication of τ and σ plus the mean, in short, $z = \tau\sigma + \mu$.

We recall from equation (8.7) above that $\lambda = W'\delta$, thus, δ is the parameter which we need to determine its form since W are variables that we conjecture will affect inefficiency. As a result, we derive the form of δ in the next section.

As we have defined earlier in equation (8.6) above, the quantity δ is part of λ. As a result, a look at equation (8.10) shows λ is not directly dependent on y but on z. On the otherhand, δ is directly dependent on λ but not on z and y. The parameter δ is only found in the second part of the distribution in equation (8.10) above. In order to find the form of δ, we would have to integrate z out of the exponential part of the distribution leaving us with the parameter δ. In order to do that we employ the Generalized Linear Model of Nelder and Wedderbun (see McCullagh and Nelder 1989).

A general representation of the generalized linear model is stated in equation (8.23) through to (8.26) as:

$$f(y_i|\theta_i, \phi) = \exp\{a^{-1}(\phi)(y(\theta_i) - b(\theta_i)) + c(y, \phi)\} \tag{8.23}$$

where y_i is a member of the exponential family of distribution, θ_i is the canonical parameter, while, $a(\cdot)$, $b(\cdot)$, and, $c(\cdot)$ are known functions.

$$\mu_i = h(\eta_i) = h(x_i'\beta) \tag{8.24}$$

$$i = 1, 2, ..., n$$

where h is referred to as the *response function* with its inverse alternative $\eta_i = g(\mu_i)$ referred to as the *link function*

Thus, for the exponential distribution:

$$f(y_i|\theta_i = \lambda_i) \equiv \prod_{i=1}^{N} f^E(y_i|\lambda_i), \qquad \lambda_i = \exp(x_i'\beta) \tag{8.25}$$

$$= \prod_{i=1}^{N} \lambda_i \cdot \exp(-\lambda_i y_i)$$

$$= \prod_{i=1}^{N} \lambda_i \cdot \prod_{i=1}^{N} \exp(-\lambda_i y_i)$$

This leads us to:

$$= \prod_{i=1}^{N} \exp(x_i'\beta) \cdot \exp\{- \sum_{i=1}^{N} \exp(x_i'\beta) \cdot y_i\} \tag{8.26}$$

Following equation (8.26), the exponential part of equation (8.10) (with focus on the paramter, δ) becomes:

$$f(\delta|\lambda) \, \alpha \, \prod_{i=1}^{N} \exp(W_i\delta) \cdot \exp\{- \sum_{i=1}^{N} \exp(W_i\delta) \cdot z_i\} \tag{8.27}$$

Leading to:

$$\alpha \, [\exp(W_i\delta)]' \cdot \iota_N \cdot \exp[- \exp(W_i\delta)' \cdot z_i] \tag{8.28}$$

Eqaution (8.28) **does not** include our proper prior, we have only gone through this process to explain to the reader the process undertaken. Thus, when we include a *normal* proper prior, we have:

$$\alpha\,[\exp{(\boldsymbol{W_i\delta})}]' \cdot \iota_{\boldsymbol{N}} \cdot \exp[-\exp{(\boldsymbol{W_i\delta})}' \cdot \boldsymbol{z_i}] \cdot [\exp\{-\tfrac{1}{2}[(\boldsymbol{\delta} - \hat{\boldsymbol{\delta}})'\boldsymbol{C}_{\hat{\delta}}^{-1}(\boldsymbol{\delta} - \hat{\boldsymbol{\delta}})]\}] \quad (8.29)$$

Equation (8.29) is the equation that we need in finding the distributional of δ. Integrating out z would give us the distributional form of δ. However, the process of integrating out z is not economical and easy. It is at this point that we need to employ the Markov chain Monte Carlo technique of the Metropolis-Hastings. Thus, in the next section we will present the simulation of δ from the Metropolis-Hastings algorithm.

The requirements of sampling using the Metropolis-Hastings method are that there must be a candidate generating density and there must be a *'probability of move'*. We view the *probability of move* like an initial rest state (which we will call *delta_bar*, $\overline{\delta}$) such that a move from this state may either be 'successful' or 'unsuccessful'. If the move is 'successful' the draw attains a new state say, '*delta*', δ. If it is unsuccessful, the draw returns into the distribution, that is, it returns into its '*delta*_bar', $\overline{\delta}$ state to try again but it is not discarded. This attribute differentiates the Metropolis-Hastings method from the *accept-reject* method (Chib and Greenberg 1995 throw more light on this subject).

We generate, *delta,* δ from the Normal density such that; $\delta = normrnd\,(\delta, \mathfrak{z})$. The standard deviation from the mean of the Normal distribution, $\mathfrak{z}$, is set to 1 (the use of one is to standardize the distribution). We define the '*probability of move'* as:

$$\varphi(\overline{\delta},\,\hat{\delta}) = min\{\frac{\pi(\delta)}{\pi(\overline{\delta})},\,1\} \qquad (8.30)$$

Alternatively, we present equation (8.30) in the logarithm form as:

$$min\{\log \pi(\delta) - \log(\overline{\delta}),\,\log(1)\} \equiv min\{logratio,\,0\} \qquad (8.31)$$

We point out that the *probability of move,*$\varphi(.)$, is actually

$$\varphi(\overline{\delta},\,\delta) = min\{\frac{\pi(\delta)\cdot q(\delta,\,\overline{\delta})}{\pi(\overline{\delta})\cdot q(\overline{\delta},\,\delta)},\,1\} \qquad (8.32)$$

$q(\overline{\delta},\,\delta)$ *and* $q(\delta,\,\overline{\delta})$ are the candidate generating densities for $\overline{\delta}$ and δ respectively which we choose from the normal distribution. Since $q(\overline{\delta},\,\delta)$ *and* $q(\delta,\,\overline{\delta})$ are from the same

symmetric distribution; they are equal and equation (8.32) reduces to equation (8.30).

Following equation (8.29), we explicitly define the density $\pi(\hat{\delta})$ as:

$$\pi(\hat{\delta}) \equiv \sum_{i=1}^{N} \exp(W_i\hat{\delta}) \cdot \exp\{-\sum_{i=1}^{N} \exp(W_i\hat{\delta}) \cdot z_i\} \cdot \exp\{-1/2 \cdot (\delta - \hat{\delta})' C_{\hat{\delta}}^{-1}(\delta - \hat{\delta})\} \quad (8.33)$$

and in logarithm form as: $\sum_{i=1}^{N} (W_i\delta) - \sum_{i=1}^{N} \exp(W_i\delta) \cdot z_i - 1/2 \cdot (\delta - \hat{\delta})' C_{\hat{\delta}}^{-1}(\delta - \hat{\delta})$

and $\pi(\bar{\delta})$ as

$$\pi(\bar{\delta}) \equiv \sum_{i=1}^{N} \exp(W_i\bar{\delta}) \cdot \exp\{-\sum_{i=1}^{N} \exp(W_i\bar{\delta}) \cdot z_i\} \cdot \exp\{-1/2 \cdot (\bar{\delta} - \hat{\delta})' C_{\hat{\delta}}^{-1}(\bar{\delta} - \hat{\delta})\} \quad (8.34)$$

and in logarithm form as: $\sum_{i=1}^{N} (W_i\bar{\delta}) - \sum_{i=1}^{N} \exp(W_i\bar{\delta}) \cdot z_i - 1/2 \cdot (\bar{\delta} - \hat{\delta})' C_{\hat{\delta}}^{-1}(\bar{\delta} - \hat{\delta})$

On inserting the logarithm forms of equations (8.33) and (8.34) into equation (8.31) we have

$$min \left\{ \left[\left(\sum_{i=1}^{N} (W_i\bar{\delta}) - \sum_{i=1}^{N} \exp(W_i\bar{\delta}) \cdot z_i - 1/2 \cdot (\bar{\delta} - \hat{\delta})' C_{\hat{\delta}}^{-1}(\bar{\delta} - \hat{\delta}) \right) \right. \right.$$
$$\left. \left. - \left(\sum_{i=1}^{N} (W_i\delta) - \sum_{i=1}^{N} \exp(W_i\delta) \cdot z_i - 1/2 \cdot (\delta - \hat{\delta})' C_{\hat{\delta}}^{-1}(\delta - \hat{\delta}) \right) \right], 0 \right\} \quad (8.35)$$

Equation (8.35) is drawn using the Metropolis–Hastings method. We utilize $\hat{\delta} = 0_p$ and $C_{\hat{\delta}} = I_p * 10$

Now, we have all it takes to carry out our simulation, however, we need to follow certain algorithms which we state as follows:

1. *set start-up value, s=0 such that $\sigma^{(s=0)}$, $\beta^{(s=0)}$, $z^{(s=0)}$, and, $\delta^{(s=0)}$*

2. *simulate a draw for $\sigma^{(s)}$ from equation (8.13) above*

3. *simulate a draw for $\beta^{(s)}$ from equation (8.18) above*

4. *simulate a draw for $\delta^{(s)}$ from equation (8.35) above such that*

$$if\ binornd(1,\ \exp(min(logratio, 1))) == 1$$

$$h = h + 1$$

$$\overline{\delta} = \hat{\delta}$$

$$end$$

5. *simulate a draw for $z^{(s)}$ from equation (8.22) above*

6. *repeat steps 2 to 6 until the values are independent of the starting values*

7. *continue to sample with g= 1, 2, ..., G*

The Metropolis-Hastings method is simple and direct; all it requires is the substitution of the derived distribution equation into the iterative algorithm. The only problem with the technique is that it is sensitive to the choice of the candidate-generating density and it requires the *tuning* of the location and scale parameters of the distribution. The Metropolis-Hastings technique also works perfectly when the distribution is in closed form and tractable. We like to mention that the Metopolis-Hastings described above is akin to the Metopolis-In-Gibbs method explained in the literature.

A total draw of G = 10,000 was made and no burn-in was used. This is because from Markov chain Monte Carlo diagnostics (the results of which we will show in the appendix), the distribution starts converging from about 1000 draws with two burn-ins. Thus, we ignored the burn-in of two and decided to run it for longer than 1000. We collect the draws such that $\{\boldsymbol{\theta}^{(g)}\}_{g=1}^{G}$ and $\{\boldsymbol{z}^{(g)}\}_{g=1}^{G}$, and then make posterior inferences based on these collected values such that our model parameter is $\boldsymbol{\theta}^{(s)} \equiv (\boldsymbol{\sigma}^{(s)}, \boldsymbol{\beta}^{(s)}, \boldsymbol{\delta}^{(s)})$ and our latent data is $\boldsymbol{z}^{(s)}$.

The techniques are a simple and easy way of obtaining 'robust' estimates for our composed error model. It relies on the use of random number generators using recognized distributions and the use of computer with high memory and processor speed (The production of computers with large memory increases the popularity of the Gibbs and Metropolis sampling, and, other MCMC techniques). The memory of the computer is an issue because personal experiences with low memory computers have shown that they are not able to deal with the demands of these samplers as they tend to give error messages or take too long to run which might cause the chain to draw from low density regions in the distribution.

Researchers are always in the quagmire of deciding the best models out of several models. In order to answer this question, they carry-out a model selection algorithm, which is the focuss the next section.

8.2.3. Model Selection/Comparison in the Composed Error Model

Hansen (2005) lists four conceptual errors - parametric vision, assuming a true data generating process, evaluating based on fit and ignoring model uncertainty - in the Frequentist method of model selection. He suggests the use of the Bayesian method of model selection in circumventing these errors. The Bayesian method of model selection is based on calculating the Bayes factor. The Bayes factor is simply, the ratio of marginal likelihoods of the different models in question. Therefore, the marginal likelihood is crucial to the calculation of the Bayes factor, as well as, model comparison. The marginal likelihood is the normalising constant of the posterior density which enables us to introduce a strict equality sign in equation (8.9). Equation (8.9) can then be re-written as:

$$\pi(\theta|y) = \frac{\sigma^{-1} \prod\limits_{i=1}^{n} f^N(y_i|x_i'\beta - z_i, \sigma) \cdot f^E(z_i|\delta)}{\prod\limits_{i=1}^{n} m(y_i)} \tag{8.36}$$

The denominator is the marginal likelihood; it is the data density without conditioning on the parameters and z. If we make the denominator the subject of equation (8.37) we have:

$$\prod\limits_{i=1}^{n} m(y_i) = \frac{\sigma^{-1} \prod\limits_{i=1}^{n} f^N(y_i|x_i'\beta - z_i, \sigma) \cdot f^E(z_i|\delta)}{\pi(\theta|y)} \tag{8.37}$$

We note that equation (8.37) is the same as that introduced by Chib (1995) which he calls the *Basic Marginal Likelihood Identity*. Thus, the basic marginal likelihood identity is the product of the data density and the prior for θ divided by the posterior for θ.

We digress to state that the marginal likelihood is actually the integral of the prior and the data density; in this case, $\theta \equiv (\sigma, \beta, \delta)$:

$$\prod\limits_{i=1}^{n} m(y_i) = \int_{-\infty}^{+\infty} \sigma^{-1} \prod\limits_{i=1}^{n} f^N(y_i|\theta) \cdot f^E(z_i|\delta) d\theta \tag{8.38}$$

We write equation (8.38) for short as;

$$m(y) = \int_{-\infty}^{+\infty} f(\boldsymbol{y}|\theta)\pi(\boldsymbol{\theta}) d\theta \tag{8.39}$$

The nature of equation (8.39) makes it difficult for econometricians to arrive at an estimate for the marginal likelihood, this has caused econometicians to develop different methods of arriving at an approximate method for its calculation. One of such methods, which is very popular among Bayesian econometricians, is the Chib (1995) which we utilize in this research. For brief explanation of other methods, we refer the reader to Didelot (2011).

Presenting equation (8.39) in its logarithm format:

$$\ln \hat{m}(\boldsymbol{y}) = \ln \ f(y|\theta^*) + \ln \ \pi(\theta^*) - \ln \hat{\pi}(\theta^*|\boldsymbol{y}) \tag{8.40}$$

The quantity $\ln \hat{m}(\boldsymbol{y})$ is the *basic marginal likelihood estimate* while $\ln \ f(y|\theta^*)$, $\ln \ \pi(\theta^*)$ and $\ln \ \hat{\pi}(\theta^*)$ are the hitherto data density, the prior and the posterior densities at the point $\theta = \theta^*$ respectively. The quantity θ^* is the parameter value at any point in the distribution. It is advisable that this point should be a high density point, which Chib (1995) describes as a maximum likelihood point in the distribution. We re-define our model parameter as $\boldsymbol{\theta}^* \equiv (\boldsymbol{\sigma}^*, \boldsymbol{\beta}^*, \boldsymbol{\delta}^*)$. Equation (8.40) is simple and straight forward, and it is not subjected to any instability problems and error, since it uses the average of the density of the values and not its inverse (Chib 1995). Also, error only enters the equation through the posterior ordinate, $\boldsymbol{\theta}^*$. This is as a result of simulation and the error due to this can be calculated (that is, the numerical standard error).

The Marginal Likelihood Estimate

From equation (8.40), it is seen that all that the calculation of the marginal likelihood requires are three quantities; the data density, the prior, and the sampling density for y (and not y and z) at the point θ^*. An appropriate posterior estimate for θ^* is available by Rao-blackwellisation such that our posterior estimate is

$$\hat{\pi}(\boldsymbol{\theta}^*|\boldsymbol{y}) = G^{-1} \sum_{g=1}^{G} \pi(\boldsymbol{\theta}^*|\boldsymbol{y}, \boldsymbol{z}^{(g)}) \tag{8.41}$$

Substituting equation (8.41) into equation (8.40), an estimate of the marginal likelihood is given as

$$\ln \hat{m}(\boldsymbol{y}) = \ln \ f(y|\theta^*) + \ln \ \pi(\theta^*) - \ln \left\{ G^{-1} \sum_{g=1}^{G} \pi\left(\theta^*|y, z^{(g)}\right) \right\} \tag{8.42}$$

The priors we utilize are the same as the ones we used in the composed error model explained earlier. We did not write a separate code for our marginal likelihood estimation,

in order words, the composed error model and the marginal likelihood estimation were all embedded in one code.

In order to obtain an estimate of the marginal likelihood, we draw for δ such that $\delta = \delta^*$ in the first run. Because δ is only directly dependent on z and not on σ, β, and y, as a result, this enables an ergodic average of the gibbs run for the posterior ordinate $\pi(\lambda|z)$ to be taken such that

$$\hat{\pi}\left(\lambda^*|z\right) = G^{-1}\sum_{g=1}^{G} \pi\left(\lambda^*|z\right) \tag{8.43}$$

The sampling continues in order to obtain the posterior ordinate of β such that $\beta \equiv \beta^*$. The parameter, β, is only dependent on σ and y and not λ and the ergodic average of its posterior ordinate is

$$\hat{\pi}(\beta^*|y) = G^{-1}\sum_{g=1}^{G} \pi\left(\beta^*|\sigma^{(g)}, z^{(g)}, y\right) \tag{8.44}$$

In the continued run we obtain the posterior ordinate for σ at the point $\sigma = \sigma^*$. Just like in the case of β, σ is dependent on y and β and not λ and its ergodic average is given as

$$\hat{\pi}\left(\sigma^*|\beta^*, y\right) = G^{-1}\sum_{g=1}^{G} \pi\left(\sigma^*|\beta^{(g)}, z^{(g)}, y\right) \tag{8.45}$$

The final step involves the substitution of these posterior estimates of θ^* above into equation (8.44) above. We now have all it takes to make posterior inference on the data. However, most econometricians would like to find the degree of variability in the estimates if the exercise is to be repeated. Chib (1995 and 2001) also provides the answer by finding the numerical standard error of the marginal likelihood estimates. He uses a combination of the delta, and, Newey and West (1987) methods to arrive at a 'robust' quantity for the standard error. Accordingly, we present the vector stochastic process thus

$$\boldsymbol{h}^{(g)} = \begin{pmatrix} h_1^{(g)} \\ h_2^{(g)} \\ h_3^{(g)} \\ h_4^{(g)} \end{pmatrix} \equiv \begin{pmatrix} \pi(y|\theta^*, z^{(g)}) \\ \alpha(\delta^{(g)}, \delta^*|y, z^{(g)}) \cdot q(\delta^{(g)}, \delta^*|y, z^{(g)}) \\ \pi(\beta^*|y, \sigma^{(g)}, z^{(g)}) \\ \pi(\sigma^*|y, \beta^{(g)}, z^{(g)}) \end{pmatrix} \tag{8.46}$$

$\pi(\boldsymbol{y}|\boldsymbol{\theta}^*, \boldsymbol{z}) = f(.|.)$ is the complete data density as

$$\hat{\boldsymbol{h}} = G^{-1} \sum_{g=1}^{G} \boldsymbol{h}^{(g)} \tag{8.47}$$

As with Markov chain processes, $\hat{\boldsymbol{h}}$ will tend towards its mean as G tends to infinity; and the variance is given by using the Newey and West (1987) method as

$$var\left(\hat{\boldsymbol{h}}\right) = G^{-1} \left[\boldsymbol{\Omega}_0 + \sum_{s=1}^{m} \left(1 - \frac{s}{m+1} \right) \left(\boldsymbol{\Omega}_s + \boldsymbol{\Omega}'_s \right) \right] \tag{8.48}$$

where

$$\Omega_s = G^{-1} \sum_{g=s+1}^{G} (\boldsymbol{h}^{(g)} - \hat{\boldsymbol{h}})(\boldsymbol{h}^{(g-s)} - \hat{\boldsymbol{h}})', \ s = 0, 1, \ldots, m$$

m is a constant value and serves as the point at which the autocorrelation function in $\boldsymbol{h}^{(g)}$ disappears. By the application of the delta method we derive the variance as:

$$var\left(\ln(\hat{m}(\boldsymbol{y}))\right) = \boldsymbol{a}' var(\hat{\boldsymbol{h}}) \boldsymbol{a} \tag{8.49}$$

Where $\boldsymbol{a} = [\hat{h}_1^{-1}, \hat{h}_2^{-1}, \hat{h}_3^{-1}, \hat{h}_4^{-1}]'$. The numerical standard error of the marginal likelihood is then given as:

$$se(\ln \hat{\boldsymbol{h}}) = \sqrt{var\left(\ln(\widehat{m(y)})\right)} \tag{8.50}$$

8.3. The Composed-Error Model Using Panel Data

In the last few sections of this chapter, we presented the composed error analysis using cross-sectional data with proper prior; apart from the reason that our presentation is pedagogic nature, it is also because two of our data sets are cross-sectional in nature. Also, because one of data set is a longitudinal in nature, we present the panel data aspect of the composed error model. The reader will notice presently that the results we obtain are similar to the earlier ones presented (we have explained the arguments in the literature on the use of improper prior in chapter five).

We set up our panel with $i = 1, 2, \ldots, n$ sample units (say household) having a one to one relationship with $t = 1, 2, \ldots, n_i$ sub-units (say time) and assuming our inefficiency is time - invariant. This implies the households' inefficiencies $z_1, z_2, \ldots z_n$ are constant over

time and z_i is scalar. We place a proper informative prior on our parameters $\theta \equiv (\sigma, \beta, \delta)$ as in the cross-sectional case and the posterior becomes

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{z}) \propto \sum_{i=1}^{n} f^N \left(\boldsymbol{y}_i | \boldsymbol{x}_i \boldsymbol{\beta} - \boldsymbol{\iota}_i \boldsymbol{z}_i, \ \sigma^2 \boldsymbol{I}_{n_i} \right) \cdot f^E(\boldsymbol{z}_i | \lambda) \tag{8.51}$$

$y_i \equiv (y_{i1}', y_{i2}', \ldots, y_{in_i}')'$ is the vector of household $i's$ production; $x_i \equiv (x_{i1}, x_{i2}, \ldots, x_{in_i k})$ is an n_i by k matrix of covariates likely to affect production while $\beta \equiv (\beta_1, \beta_2, \ldots, \beta_k)'$ is $k \times 1$ vector of location and λ is as defined in equation (8.6). Equation (8.51) is the same as equation (8.9) except that we have written the former in a less explicit case and we introduced time into it.

Except for a few changes, the conditional distributions of our model parameters are same as in the cross-sectional situation. The linear regression scale parameter,σ, is of the same conditional distribution as in the cross-sectional case. In other words, it is of the inverse–Gamma distribution;

$\pi(\sigma|\boldsymbol{\beta}, \boldsymbol{z}, \boldsymbol{y}) \equiv f^{IG}(\sigma|v, \ s^2)$, $v = n$, $vs^2 = (\boldsymbol{y} - \boldsymbol{x}\boldsymbol{\beta} + \acute{\boldsymbol{w}}\boldsymbol{z})'(\boldsymbol{y} - \boldsymbol{x}\boldsymbol{\beta} + \acute{\boldsymbol{w}}\boldsymbol{z})$, $y = (y_1', y_2', \ldots, y_n')'$, $x = (x_1', x_2', \ldots, x_n')$ and $\acute{\boldsymbol{w}} = diag(\boldsymbol{\iota}_1, \boldsymbol{\iota}_2, \ldots, \boldsymbol{\iota}_n)$ is an S x N matrix with one to one interaction with $\boldsymbol{z}$; $u = (u_1', u_2', \ldots, u_3')'$. Likewise the frontier parameter, β, has a multivariate-Normal distribution; $\pi(\boldsymbol{\beta}|\sigma, \boldsymbol{z}, \boldsymbol{y}) \equiv f^{mN}(\beta|\hat{\boldsymbol{\beta}}, \boldsymbol{C}_{\hat{\beta}})$, $\hat{\boldsymbol{\beta}} = \boldsymbol{C}_{\hat{\beta}}^{-1} \boldsymbol{x}'(\boldsymbol{y} + \acute{\boldsymbol{w}}\boldsymbol{z})$, $\boldsymbol{C}_{\hat{\beta}} = \sigma^2(\boldsymbol{x}'\boldsymbol{x})^{-1}$. The vector of the inefficiency term is truncated Normal and the truncation is at zero, $\pi(\boldsymbol{z}|\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\omega}, \boldsymbol{y}) = f^{TN}(\boldsymbol{z}|\hat{\boldsymbol{z}}, \boldsymbol{C}_{\hat{z}})$, $\hat{\boldsymbol{z}} = \boldsymbol{C}_{\hat{z}}^{-1}(\boldsymbol{c}'\boldsymbol{h}\lambda - \acute{\boldsymbol{w}}'\boldsymbol{b}\boldsymbol{a})$, $\boldsymbol{C}_{\hat{z}} = (\acute{\boldsymbol{w}}'\boldsymbol{b}\acute{\boldsymbol{w}} + \boldsymbol{c})$, $\boldsymbol{a} = \boldsymbol{y} - \boldsymbol{x}\boldsymbol{\beta}$, $\boldsymbol{b} = \sigma^{-2}\boldsymbol{I}_s$, $\boldsymbol{c} = \boldsymbol{\omega}^{-2}\boldsymbol{I}_n$ and $\boldsymbol{h}$ is the n-dimensional unit vector. The form of the derived distributions for δ are the same as the form which we derived earlier in section 8.2.2 above. Similarly, the calculation of the marginal likelihood and the numerical standard $h_n^{(g)}$ follow exactly the same steps as described above in section 8.2.3.

So far, this section succeeds in providing the framework for our analysis, in the next section, we discuss the summary of this chapter.

8.4. Summary

We have been able to input our theoretical framework into econometric model, which led us to the pedagogic presentation of the exponential model forms of the composed error models. Therein, we explained how the individual parameter distributions are arrived at and how they are drawn using the Makov chain Monte Carlo techniques of Gibbs and Metropolis-Hastings sampling. We also took a journey into marginal likelihood calculation which is essential for the Bayesian method of model comparison.

We have assumed the exponential distribution for our inefficiency values, however, we would like to state that there are other distributions that a researcher can assume for the

inefficiency variable. These distributions include the truncated-normal, gamma, and, the half-normal distributions. The exponential distribution as we have seen places restriction on the inefficiency values. Apart from this, it is a one parameter distribution with fixed shape parameter (*see appendix for sample codes on different matlab© exercises*). Next, we present the results of our Markov chain Monte Carlo diagnostics and estimations from the use of the techniques described above.

9. Empirical Results

9.1. Introduction

In Chapter eight, we presented the empirical framework of this research. In this chapter, we present the results of our analyses. As a precursor to a detailed presentation of the results of the composed-error (stochastic frontier) model, we carried out a covariate selection test, as well as, a model selection test on each of the selected countries (two Markov chain Monte Carlo diagnostics were carried out on the data the results of which are presented in the appendix). The models considered for selection were the normal linear regression model, the composed error model with malaria on the frontier part, the composed error model with malaria on the inefficiency part, and, composed error model with malaria on both parts of the model.

We would like to emphasize that the whole of this chapter is focused on equation (4.13) which is the centre of our research. We recall that this equation consists of the price (p_a) which multiplies the malaria estimate from the composed error model and a constant value set by the researcher. The result of this multiplication gives us a measure of the willingness to pay for malaria prophylactic measures. We present both the point estimate and the posterior distribution for each of the countries.

We start off by presenting the results of the covariate selection exercises in section 9.2, and, render a pedagogic presentation of the results of the stochastic frontier analyses on each of the selected countries in section 9.3. We discuss the results in detail and make inferences about them.

9.2. The Covariate/Model Selection Results

In this section, we present the result of the covariate selection algorithm for each of the countries in our research. The covariate selection algorithm uses the marginal likelihood calculation discussed in the last chapter. We undertake two phases in our covariate selection exercise. The first phase involves determining if a particular variable should be included in our model or not, this is regardless of whatever model we intend to analyse?

The second phase, which is a model selection exercise, determines what type of model the researcher should analyse. We considered the normal linear regression model, the composed error model with malaria on the frontier part, the composed error model with malaria on the inefficiency part, or the composed error model with malaria on both parts

of the model. In other words, this exercise helps to decide which (frontier or inefficiency) part of the model a particular variable goes. This remedies the argument that "on which part of the composed-error model should the malaria variable go"? Hence, we present two different covariate/model selection results.

We note that the second phase compliments the first, in other words, if a variable is selected in the first phase, it is expected to have a high probability in the second phase. An exception to the rule occurs when a variable is not selected in the first phase, but has a high probability in the second phase. If this scenario happens, just as will be seen in the Nigeria case presently, the researcher may decide to include this variable in his analysis. We start off with the presentation of the results of the Nigeria data, followed by the Ethiopia data, and then the Tanzania data.

9.2.1. The Nigeria Data Covariate/Model Selection Result

In the Nigeria data, we ran the algorithm on seventeen variables which are labour, wage, the number of days worked on the farm, sales, size of land cultivated, the amount of seeds used, age, if the land was irrigated or not, pesticide use, equipment use, fertiliser usage, use of herbicide, use of animal traction, gender, education, malaria prevalence, and, constant. We present the result of the covariate selection phase in figures 9.1 and 9.2 below:
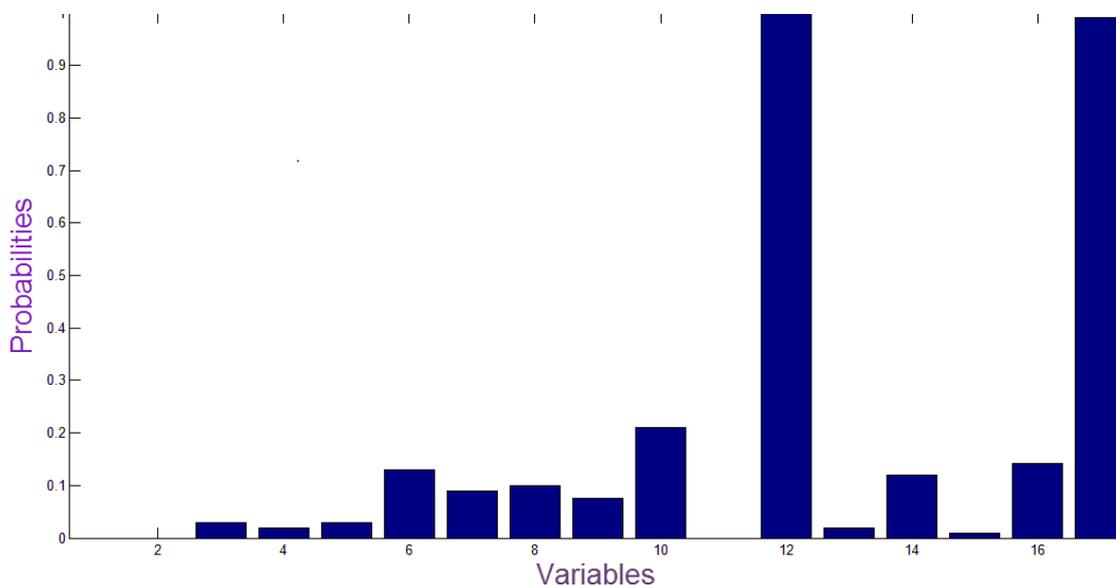
Figure 9.1.: Covariate Selection Result for the Nigeria Data

1=labour, 2=wage, 3=number of days spent on the farm, 4=land cultivated, 5=seed used, 6=age, 7=total sales, 8=irrigation, 9=pesticides, 10=equipment used, 11=fertiliser use, 12=herbicide use, 13=animal traction, 14=gender, 15=education, 16=malaria prevalence, 17=constant.
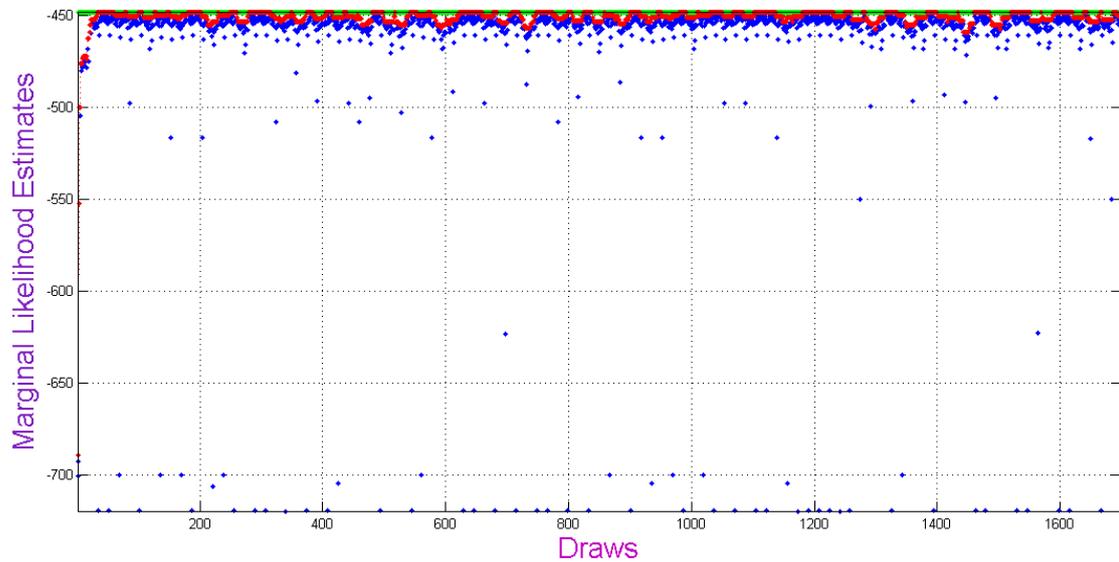
Figure 9.2.: Marginal Likelihood Estimates of the Nigeria Data Covariate Selection
Model
The red line is the accepted probability density, while the blue dots are proposal probability densities.

In this case, the criteria for selection is any probability value greater than 0.02. Setting
the criteria is at the researcher's discretion as this is based on the number of variables with
high probability scores. If the probability values are low (as in this case), the researcher
should lower the criteria to accommodate as many variables as reasonably possible and
vice-versa.

From table 9.1, the selected variables are: *land cultivated, age, total sales, pesticides,
equipment used, herbicide, animal traction, gender, malaria prevalence, and, the constant
term.* However, as will be seen shortly, we selected the labour variable because it gave a
high probability value in the second phase.

Next, we present the result of the model selection exercise. We display the result of the
composed-error model with the malaria variable on the frontier part of the equation, this
is because, it has the highest marginal likelihood estimate (of -280) of the four models in
the second phase of our analysis.
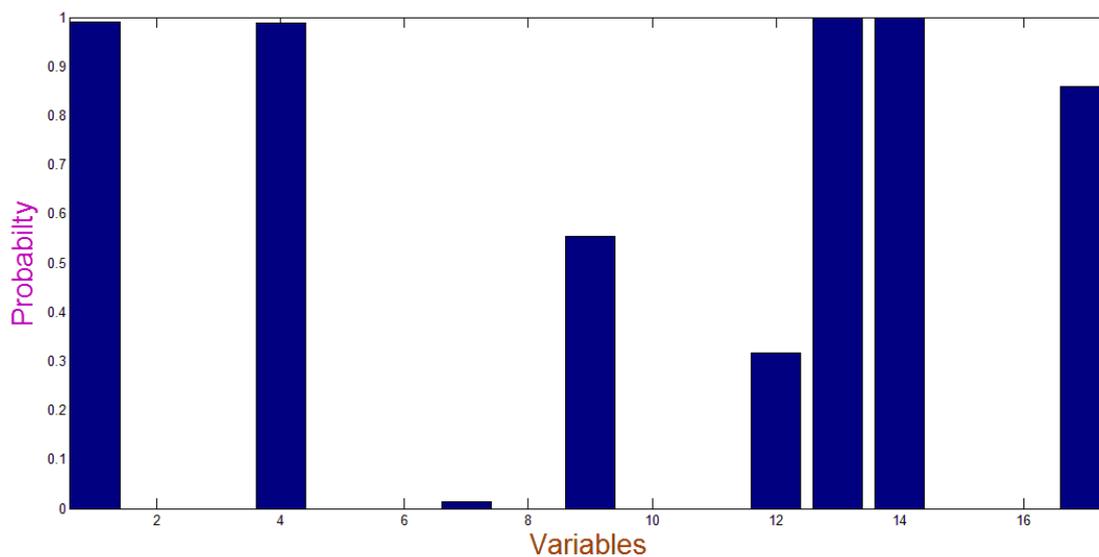
Figure 9.3.:   The Nigeria Model Selection Result Showing Variables that Should be on the *Inefficiency* Part of the Composed-Error Model

1=labour, 2=wage, 3=number of days spent on the farm, 4=land cultivated, 5=seed used, 6=age, 7=total sales, 8=irrigation, 9=pesticides, 10=equipment used, 11=fertiliser use, 12=herbicide use, 13=animal traction, 14=gender, 15=education, 16=malaria prevalence, 17=constant.
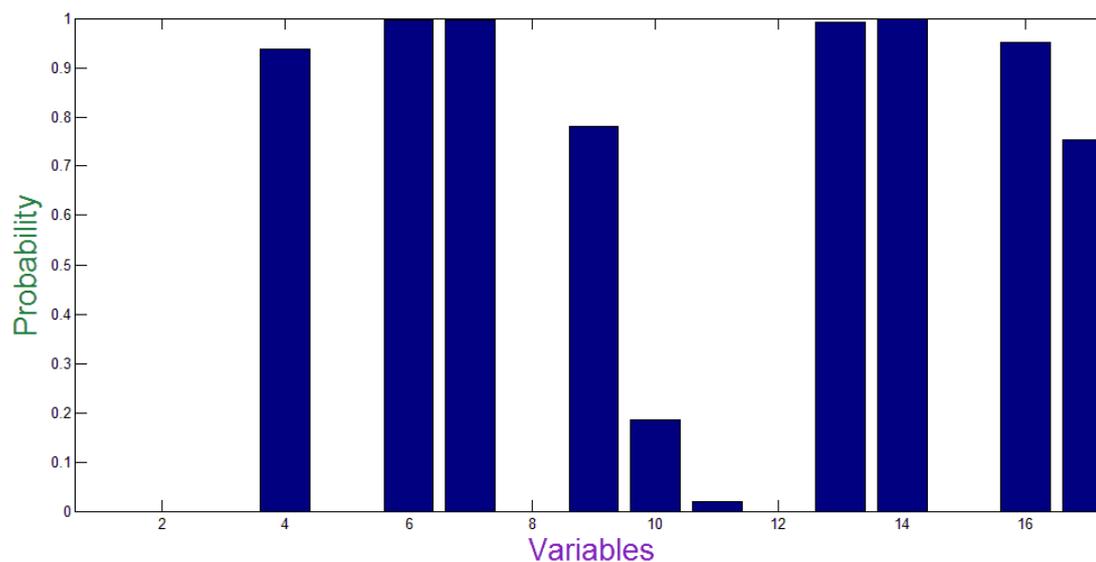
Figure 9.4.:   The Nigeria Model Selection Result Showing Variables that should be on the *Frontier* Part of the Composed-Error Model

1=labour, 2=wage, 3=number of days spent on the farm, 4=land cultivated, 5=seed used, 6=age, 7=total sales, 8=irrigation, 9=pesticides, 10=equipment used, 11=fertiliser use, 12=herbicide use, 13=animal traction, 14=gender, 15=education, 16=malaria prevalence, 17=constant.
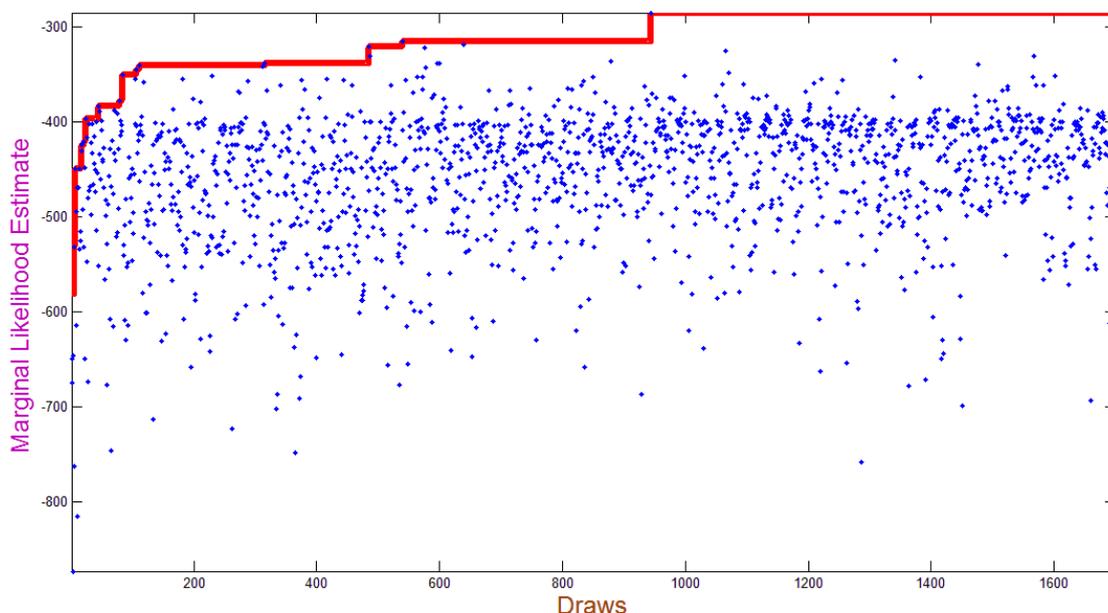
Figure 9.5.: Marginal Likelihood Estimate of the Model with *Malaria* on the Frontier part of the Composed-Error Model

This model has a marginal likelihood estimate of -280 which is the highest of the three models in the second phase of the model selection exercise.

Figures 9.3 and 9.4 show some variables are on both sides of the model, in this situation, we chose the side with the highest probability except for the "gender" covariate with same probability values in both figures. In this case, the variable appears on both sides of the model. We include the variable "sales" in our analysis to emphasize its importance.

On looking at the results of the covariate/model selection exercises, the covariates we use for the composed error analysis are as follows:

The Frontier part: Age, pesticides, equipment used, gender, malaria, and, the constant.

The Inefficiency part: labour (this is included because it has a large probability in the model selection exercise), land, herbicide, gender, animal traction, and constant.

9.2.2. The Ethiopia Data Covariate/Model Selection Result

We ran the algorithm on thirteen variables which are: land area cultivated, labour, fertiliser, land type, index of other input, asset, credit, slope, intercropping, age , education, malaria and the constant. We commence by presenting the covariate selection result from the first phase of the exercise below:

179
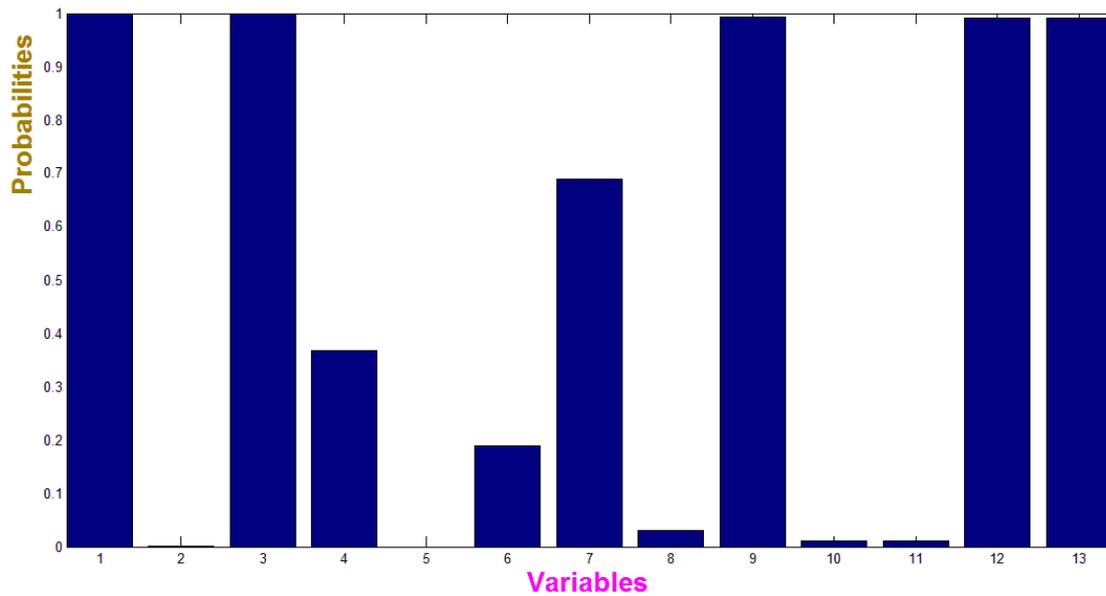
Figure 9.6.: Covariate Selection for the Ethiopia Data

1.landarea 2. labour 3. fertiliser 4. landtype 5. otherinput 6. asset 7. credit 8. slope 9. intercropping 10. age 11. education 12. malaria 13. constant
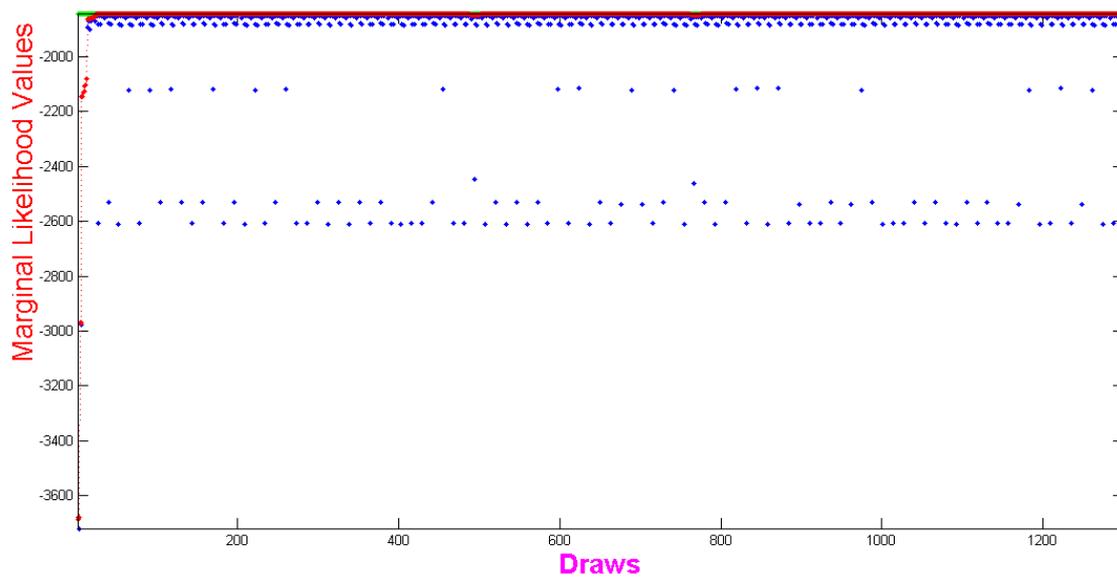
Figure 9.7.: Marginal Likelihood Estimates of the Nigeria Data Covariate Selection Model

Red lines are the accepted densities and blue lines are the proposal densities

Using a criteria of ≥ 0.05, the selected variables are: *land area, fertilizer, land type, asset, credit, slope, intercropping, malaria, constant.*

Next, we present the result of the model selection exercise. We display the result of the composed-error model with the malaria variable on the inefficiency part of the model, this is because, it has the highest marginal likelihood estimate (of 1406) of the four models in the second phase of our analysis.
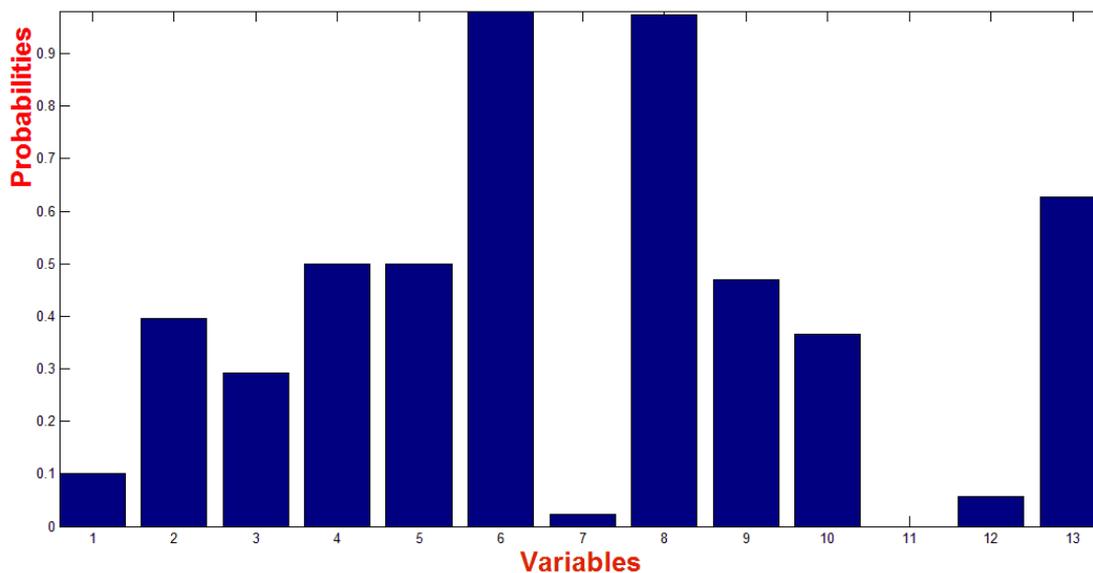
Figure 9.8.:   The Ethiopia Model Selection Result for *frontier* Part of the Composed-Error Model

1.landarea 2. labour 3. fertiliser 4. landtype 5. otherinput 6. asset 7. credit 8. slope 9. intercropping 10. age 11. education 12. malaria 13. constant
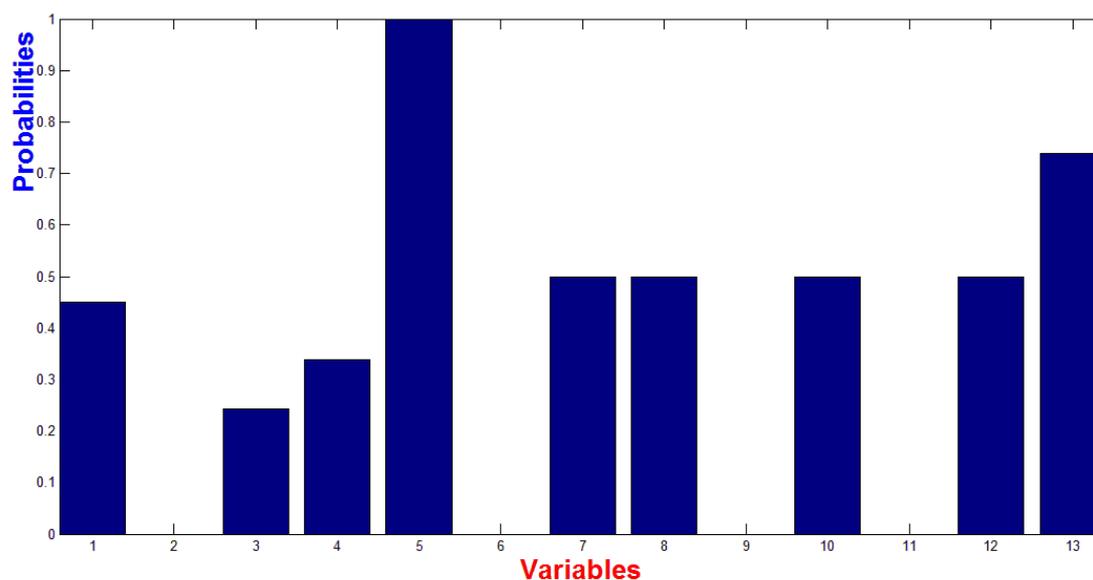
Figure 9.9.:   The Ethiopia Model Selection Result for *inefficiency* Part of the Composed-Error Model

1.landarea 2. labour 3. fertiliser 4. landtype 5. otherinput 6. asset 7. credit 8. slope 9. intercropping 10. age 11. education 12. malaria 13. constant
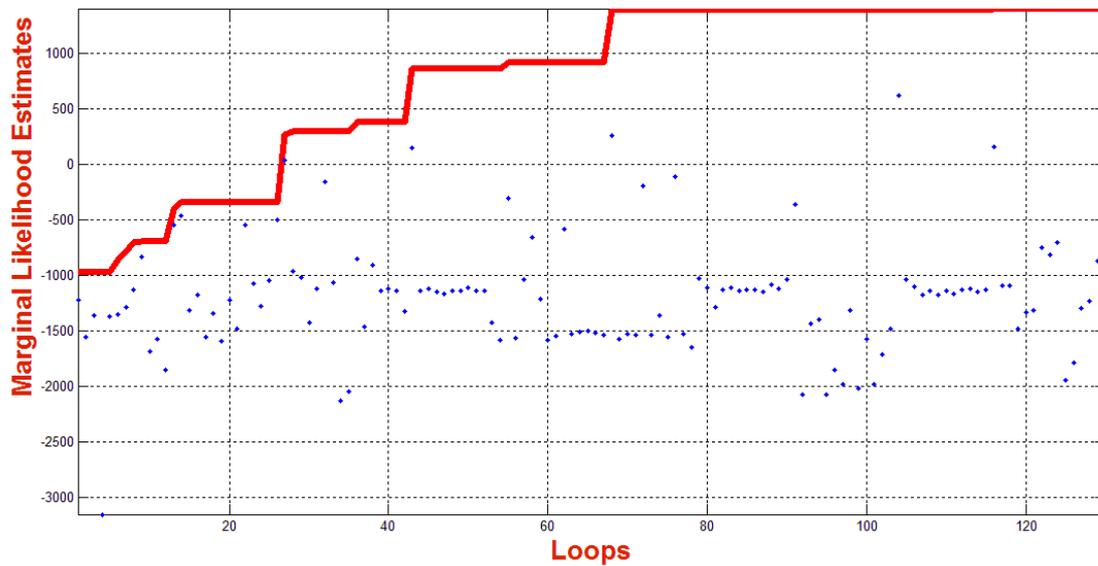
Figure 9.10.: Marginal Likelihood Estimate of the Model with *Malaria* on the inefficiency part of the Composed-Error Model
*We produce the mean of ten loops

On looking at 9.8 and 9.9 it could be seen that the malaria variable on the inefficiency part of the model has the highest probability, thus, we put malaria on the inefficiency part of the model.

9.2.3. The Tanzania Data Covariate/Model Selection Result

In the Tanzania data, we utilize thirteen variables, these are: *land area cultivated, total labour used on the farm, index of other inputs used on the farm, use of erosion control technology, the soil quality, mean of education, access to extension services, access to credit, age, total sales, the malaria prevalence and the constant.* The figure below presents the result of the first phase, which is the covariate selection algorithm:
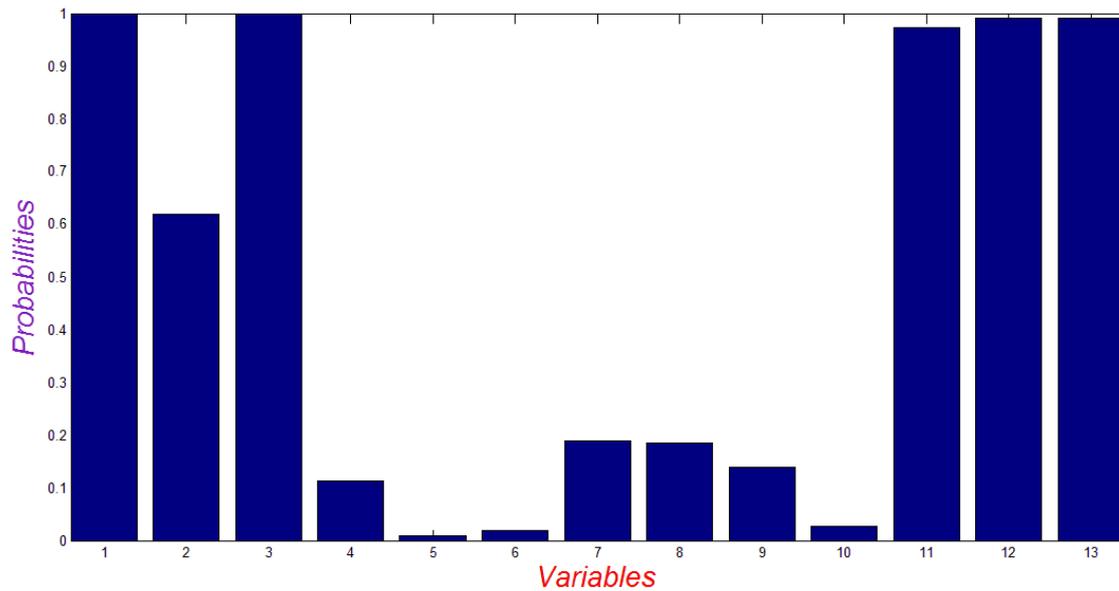
Figure 9.11.: Covariate Selection Result for the Tanzania Data

1= logarea, 2 = loglabour, 3 = logotherinputs, 4 = erosion control, 5 = soilquality, 6 = meane ducation, 7= extension, 8 = inputcredit, 9 = age, 10 = intercropping, 11 = logsales, 12 = malariaprevalence, 13 = constant
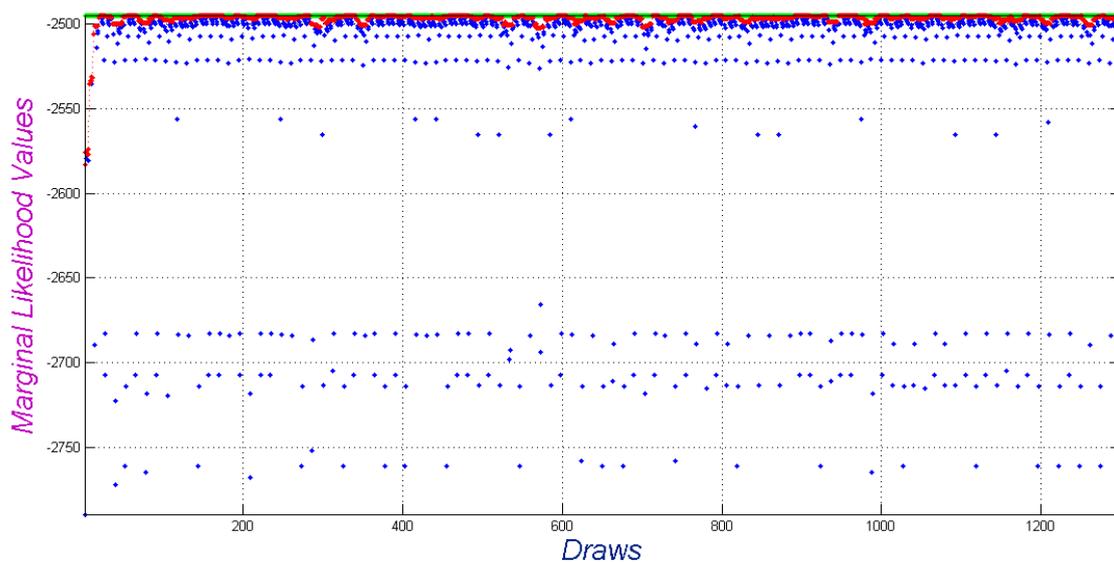
Figure 9.12.: Marginal Likelihood Estimates of the Tanzania Data Covariate Selection Model

The red line is the accepted probability density, while the blue dots are proposal probability densities.

We selected variables with probabilities greater than or equal to 0.05. Thus, the variables selected were; *land area cultivated, total labour, index of other inputs, use of erosion control technology, access to extension services, access to credit, age, practise of intercropping, total sales, malaria prevalence, and, the constant.*

The next exercise is the model selection exercise. Here, we decide on which side of the composed error model our malaria variable should be in. As stated earlier, the four models considered are the normal linear model, the composed error model with malaria on the

frontier side; the composed error model with malaria on the inefficiency side of the model, and, the composed error model with malaria on both sides of the model. We report the model with malaria on the frontier part of the model, this is because, its marginal likelihood value was higher by about seventy points. It had a marginal likelihood of 1981 (see figure 9.15 below).
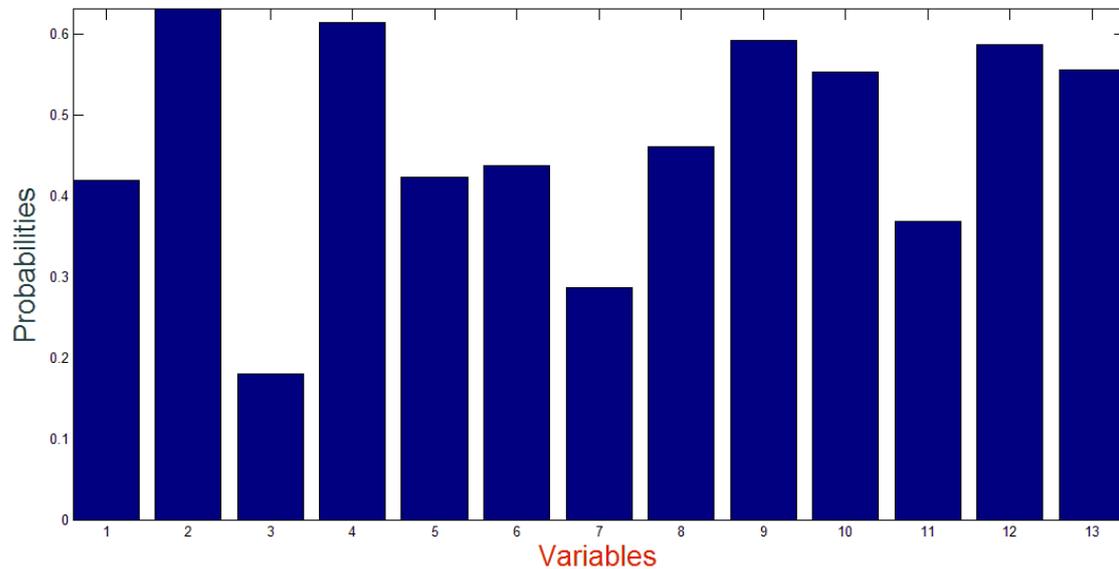
Figure 9.13.: The Tanzania Model Selection Result Showing Variables that should be on the *Frontier* Part of the Composed-Error Model

1= log of area cultivated, 2 = log of labour used, 3 = index of otherinputs, 4 = erosion control, 5 = soil quality, 6 = mean education, 7 = extension, 8 = credit, 9 = age, 10 = intercropping, 11= log of total sales, 12 = malaria prevalence, 13 = constant
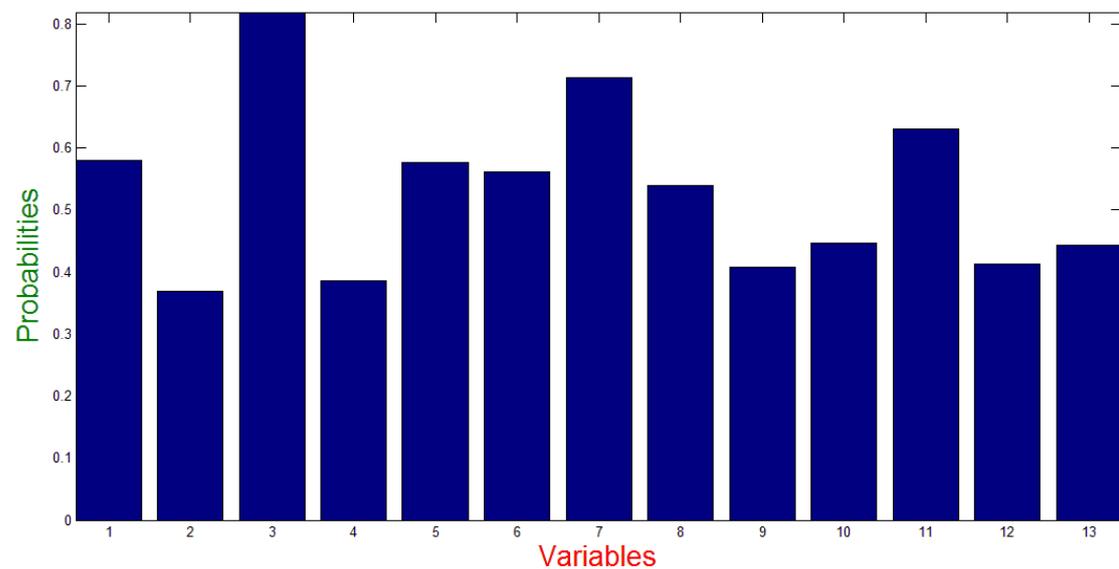
Figure 9.14.: The Tanzania Model Selection Result Showing Variables that should be on the *Inefficiency* Part of the Composed-Error Model

1= log of area cultivated, 2 = log of labour used, 3 = index of otherinputs, 4 = erosion control, 5 = soil quality, 6 = mean education, 7 = extension, 8 = credit, 9 = age, 10 = intercropping, 11= log of total sales, 12 = malaria prevalence, 13 = constan
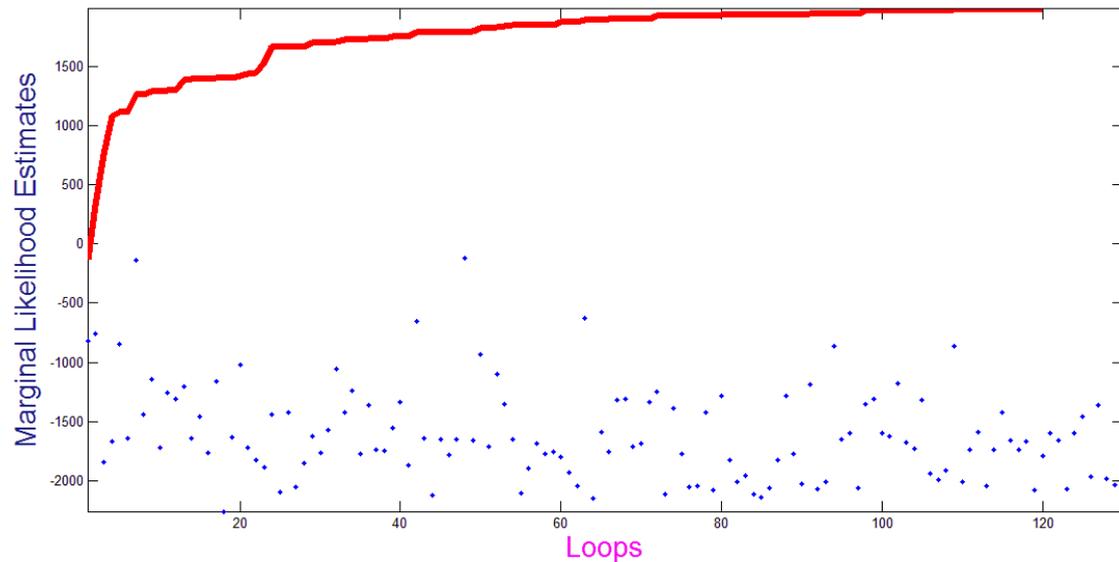
Figure 9.15.: Marginal Likelihood Estimate of the Model with *Malaria* on the Frontier part of the Composed-Error Model

*The x-axis is the mean value obtained after running the chain for ten additional loops

Comparing figures 9.13 and 9.14, it could be seen that the malaria variable in the frontier part of the model has higher probability values than that in the inefficiency part of the model. As a result, we will be analysis our model with the malaria variable on the frontier part of the model.

9.3. The Composed Error Model Results

The last section presents the covariate/model selection results. This section reports the results of the composed-error model from the three countries. We present the results in tables 7.1, 7.2, and 7.3 and discuss our findings. Tables 7.1, 7.2, and 7.3 show the Nigeria, Ethiopia, and Tanzania data respectively. Each of these tables show a willingness to pay value which is the product of the price of agricultural staples and the malaria prevalence estimate (and a constant which we set at one). In order to show the level of uncertainty associated with this calculation we present the posterior distribution.

We digress to remind the reader that malaria prevalence is measured in *case per 1,000 individuals per annum* (Hay et al. 2009). For example, a standard adopted for malaria stable areas is ≥ 0.1 case per 1,000 individuals per annum.

9.3.1. Nigeria

Table 9.1.: Estimates of the Composed Error Model

S/N	Variables	Nigeria	
		Mean	95% Highest Posterior Density
	frontier Variables		
1.	age	0.004	[-0.14 0.15]
2.	pesticides	-0.020	[-0.18 0.15]
3.	equipment_use	0.120	[-0.02 -0.27]
4.	malaria	0.212	[0.05 0.37]
5.	constant	0.028	[-0.12 0.18]
6.			
	Inefficiency Variables		
10.	labour	-0.062	[-0.24 0.19]
11.	land	-0.028	[-0.23 0.08]
12.	herbicide	-0.048	[-0.21 -0.20]
13.	animal_traction	0.180	[-0.04 0.34]
14.	gender	-0.087	[-0.34 0.07]
15.	constant	0.170	[-0.09 0.27]
16.			
17.			
21	wtpay	0.10	N/A
22	log.max.likeli.	-402.85	N/A
23	log.marg.likeli.	-402.20	N/A
24	nse	0.000	N/A
25	R^2	0.052	N/A
26	Average eff.score	32.85	N/A
27	mean predic. val.	-0.45	N/A
28			

*significant at 95% highest posterior density interval/level of signicance

Figure 9.16.: Willingness-To-Pay Posterior Distribution (Values in Nigeria Naira)
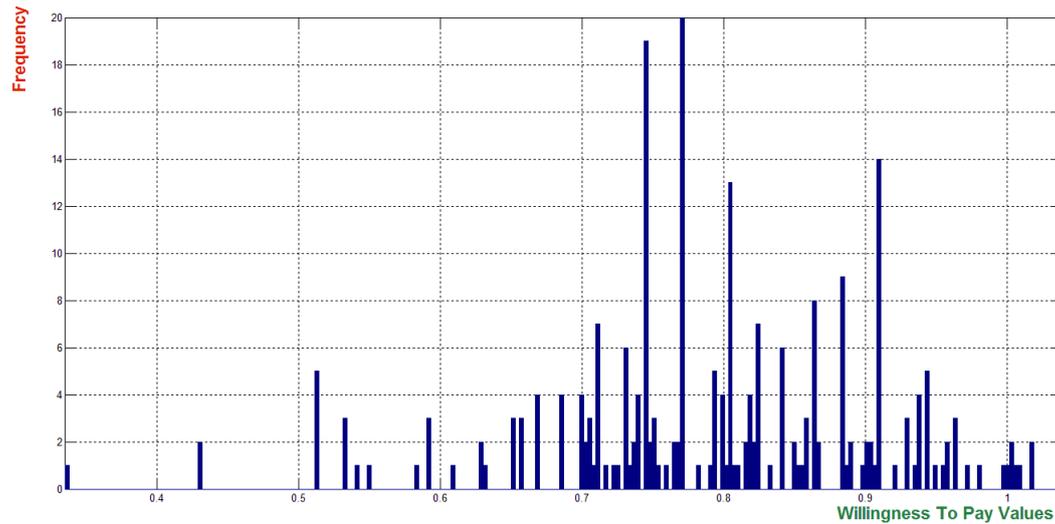
A look at the frontier part of the model shows that all the variables (including malaria prevalence) have the positive sign except for the pesticide variable. We cannot explain why the malaria variable does not have the expected sign (the expected sign is negative) but one reason we may adduce for this is the fact that the malaria data was not collected by the same organisation, thus, the values used do not bear direct relation to each household in our data. The pesticide variable does have the expected sign (it has the negative sign). This might be a situation of overuse of pesticides by the farmers or they might have used the wrong pesticides for their crops, such that, its application results in productivity decline.

On the inefficiency part of the model, all the variables except for the animal traction variable (and constant) have the negative sign. Thus, an increase in all of these variables is expected to cause inefficiency to decrease, except for the animal traction variable that will cause inefficiency to increase. The use of other equipment in place of animals might be a reason for this.

All the variables in our model are not significant except for the malaria prevalence variable. Consequently, a 21 percent increase in malaria case per 1000 per annum will cause productivity to increase by about 100 percent.

On multiplying the price of staples by the malaria prevalence estimate, our result shows that for a 100 percent increase in malaria case per 1000 per annum, farmers are willing to pay an average amount of about 10 Nigerian Naira (about US$0.1 at the 2009 conversion rate) for malaria preventive drugs. The result of the posterior distribution in table 7.16 above is negatively skewed. It shows that most farmers in the sample are willing to pay between 70 Nigerian Naira (US$0.47) and 93 Nigerian Naira (US$0.63) for a 100 percent increase in malaria case per 1000 per annum.

The mean efficiency score is about 30%, which shows that the farmers in our sample

are highly inefficient. This therefore negates Schultz's hypothesis that peasant farmers are poor but efficient. The log marginal likelihood is about minus four hundred. The regression coefficient has a poor goodness of fit value of 5%.

9.3.2. Ethiopia

Table 9.2.: Ethiopia Estimates of the Composed Error Model

S/N	Variables	Ethiopia	
		Mean	*95% Highest Posterior Density*
frontier Variables			
1	fertiliser	1.44*	[0.97 1.92]
2	asset	0.50	[-0.29 1.33]
3	slope	-0.44*	[-0.79 -1.08]
4	intercropping	-0.73*	[-1.061 -0.406]
5	constant	7.13*	[6.65 7.62]
inefficiency Variables			
6	land area	3.31*	[2.56 3.86]
7	fertiliser	2.41*	[1.46 3.19]
8	credit	0.45*	[0.22 0.64]
9	malaria	0.81*	[0.47 1.11]
10	constant	-0.78*	[-1.16 -0.53]
21	wtpay	0.11	N/A
25	R^2	0.79	N/A
26	Average eff.score	0.89	N/A
27	mean predic. val.	6.14	N/A

*significant at 95% highest posterior density interval/level of signicance

Figure 9.17.: Ethiopia Willingness-To-Pay Posterior Distribution (Values in Ethiopia Birr)
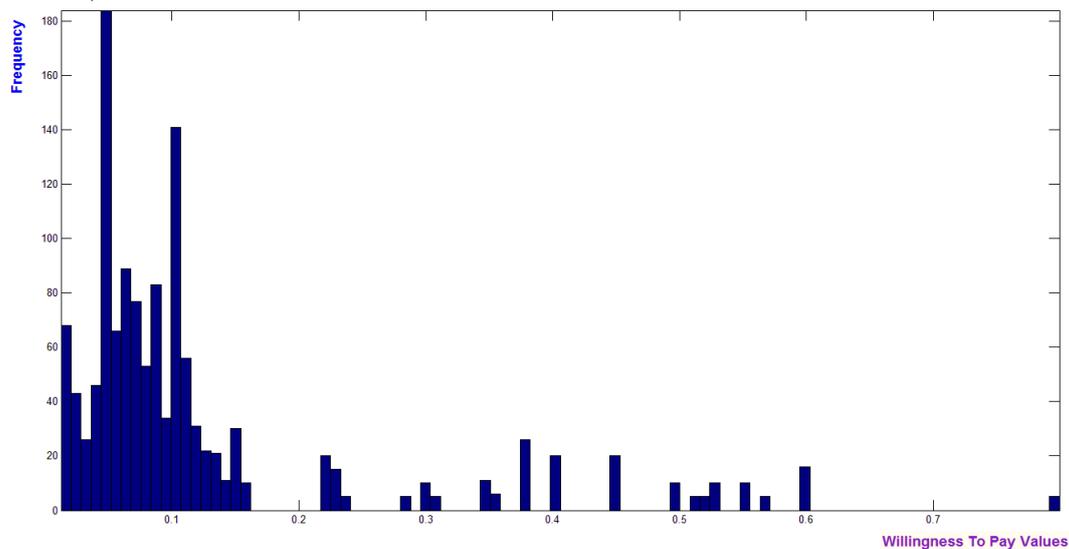
The frontier part of the table shows that apart from asset, all the variables are significant at 5%. Thus, they significantly affect productivity. Fertiliser and asset are positive thus an increase in these variables will cause an increase in productivity, other things being equal and *vice versa*. Intercroppping and slope are negatively related to productivity. Thus, any increase in any of these variables will cause productivity to decline. The result for intercropping is quite surprising because it is expected to increase productivity. This may be because the farmers plant on the same land similar crops that require similar soil nutrients.

All of our variables on the inefficiency part of the model are significant including the malaria variable at 5%. Apart from the constant term, all the variables are positive. The sign for our malaria variable follows the expected outcome (although the signs for the other variables did not follow the expected outcome as we expect them to be negatively related to inefficiency). Our result shows that 81 percent increase in malaria case per 1000 per annum will cause a 100 percent increase in inefficiency and *vice versa*.

On multiplying the price of staples by the malaria prevalence estimate, we obtain a willingness to pay value of about 11 Ethiopian birr (about US$0.98 at 2009 conversion rate) if malaria increases by 100 percent case per 1000 individuals per annum. The posterior distribution in figure (7.17) is positively skewed and shows that most farmers in the distribution are willing to pay between 0 Birr (US$0) and about 15 Ethiopian Birr (US$1.34) if malaria increases by 100 percent case per 1000 individuals per annum.. The analysis shows a high regression coefficient value of 79%. This means our model fits the data very well. The mean predicted value is about 6.14 kilograms per hectare.

Our results reveal that Ethiopian farmers have a mean efficiency score of about 90%. This assertion supports Schultz's hypothesis, that peasant farmers are poor but efficient.

9.3.3. Tanzania

Table 9.3.: Tanzania Estimates of the Composed Error Model

S/N	Variables	Tanzania	
		Mean	*95% Highest Posterior Density*
	frontier Variables		
1	labour	0.86*	[0.24 1.44]
2	erosioncontrol	-0.11	[-0.29 0.07]
3	credit	0.04	[-0.63 0.71]
4	age	-0.87	[-1.42 -0.15]
5	malariaprevalence	-0.75	[-1.01 -0.47]
6	constant	6.75*	[3.90 9.60]
	inefficiency Variables		
7	areaplanted	-0.21	[-0.49 0.04]
8	otherinputs	0.14	[-0.60 0.68]
9	extension	0.32	[-0.48 1.18]
10	malariaprevalence	0.16	[-0.25 0.50]
11	constant	-0.49*	[-0.86 -0.04
12	R^2	0.76	n/a
13	Average eff.score	0.86	n/a
14	wtp	0.96	n/a
15	mean predic. value	6.22	n/a

*significant at 95% highest posterior density interval/level of signicance

Figure 9.18.: Posterior Distribution (Values in Tanzanian Shilling)
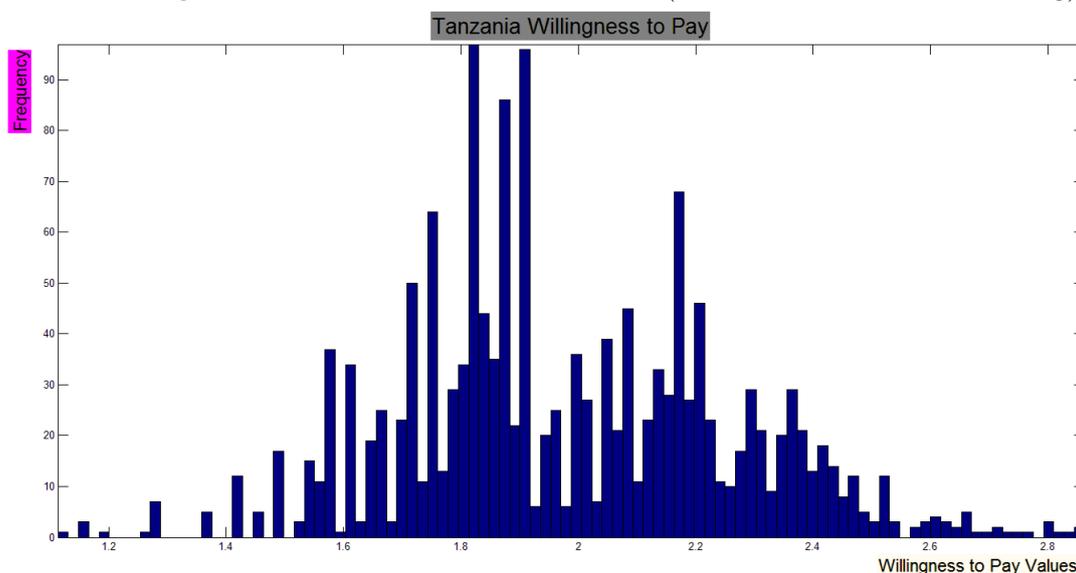
Apart from the constant term, only labour is significant at 5% on the frontier part of the model, and, thus it significantly affects productivity. It has a positive value, hence, a 100 percent increase in labour will cause productivity to increase by 86 percent. The

credit and constant variables were also positive, while, age, and malaria prevalence are negative. The negative sign on the malaria prevalence variable follows initial expectation as a decrease in malaria prevalence is expected to cause productivity to increase and *vice-versa.*

No variables were significant in the inefficiency part of the model. Only "area planted" and the "constant" variables are negative. All other variables are positive. The malaria prevalence variable on the inefficiency side of the follows the initial expectation as an increase in malaria prevalence is expected to cause inefficiency to increase.

On multiplying the price of staples in the data by the malaria point estimate, we obtain a willingness to pay value of about ninety-six Tanzania-Shillings (about US$0.07 at 2009 conversion rate) if there is a 100 percent increase in malaria case per 1000 individuals per annum. The posterior distribution in table 7.18 above is slightly right skewed and it shows a modal willingness to pay value of 180 Tanzanian-Shillings (about US$0.13) if there is a 100 percent increase in malaria case per 1000 individuals per annum.

The average efficiency score is about ninety percent. This assertion supports Schultz's hypothesis, that peasant farmers are poor but efficient. There is a mean prediction value of six kilograms per hectare. The goodness of fit of our model is about eighty percent.

9.4. Summary

In this chapter, we attempted to investigate the significance of the malaria in productivity and efficiency measurement in three African countries; Nigeria, Ethiopia, and, Tanzania.

In chapter 4, we suggested that malaria should influence the farmers' efficiency more, but the results of the covariate/model selection exercises for Nigeria and Tanzania suggest otherwise. In Nigeria, it states that the malaria variable should be on the frontier part of the model, while in Tanzania, it states that it should be on both parts of the model. However, the result from the analysis of the Ethiopian data supports our position. Thus, these findings support the argument in the literature that there is no universally accepted rule that certain variables should be on certain part of the composed error model. Our use of the Bayesian covariate selection algorithm and the Markov chain diagnostics is a promising way of reaching a universal consensus.

We could not exactly explain why the malaria variable has the positive sign in the Nigeria data this might be due to measurement error. This is because the malaria covariate values used in this research did not measure the malaria prevalence in each household rather the malaria values used only measure the prevalence in the area where the households are located. From a Bayesian perspective, this does not mean that our method is wrong this is because the model selection exercise supersedes the sign and significance of variables. The malaria variables in Ethiopia and Tanzania follow the expected sign.

On the average, farmers in all our samples are willing to pay less than one United States dollar for a 100 percent increase in malaria case per 1000 individuals per annum. The posterior distribution of the three countries shows that farmers in Ethiopia are willing to pay more for malaria abatement than farmers in Nigeria and Tanzania.

In the next chapter, we present the conclusions of our research.

10. Summary And Concluding Remarks

The last chapter presents the results of our analysis. In this chapter, we provide the synopsis of our research and then provide relevant remarks about it. Thus, section 10.1 furnishes the reader with a summary of this research, we then provide the likely policy implications of this work in section 10.2 and also highlight the merits and limitations of this study in sections 10.3 and 10.4 respectively.

10.1. Summary

Malaria is one of the diseases that plagues the African continent, as a result, it is a key health issue in Africa. A lot of malaria control strategies (these include the anti-malaria drugs, mosquito nets and its variants, use of insecticides) have been put in place or suggested by policy makers, but one problem that both policy makers and researchers are always interested in; is a measure of the amount the farmer is <u>willing to pay</u> for any malaria control measures. The literature (for example, Onwujekwe et al, 2004 and 2005, Jimoh et al. 2007) attempts to use the non-market based method of stated preference to arrive at a measure of the willingness of the farmer to pay for malaria abatement. The subjective nature of the approach, the epidemiology of the disease, the nature of the control measures, and, the paucity of accurate malaria incidence data in most African countries makes their measure their inexact.

Thus, our study attempted to develop and introduce a method of arriving at a good measure of the amount that the household is willing to pay for malaria prevention in three African countries - Nigeria, Ethiopia, and Tanzania. In doing this, we employed two fundamental models that have far-reaching foundation in the literature on the study of the rural household - the household production and the composed-error models (we scrutinized the literature on these two models). During the course of arriving at this quantity, we also provided a pedagogic development of the Bayesian hierarchical approach to the exponential composed-error model.

To arrive at this measure, we have employed two different types of data sets for each of our country of study - the household and the spatial malaria prevalence data sets. Our research has also highlighted how to obtain numerical values for malaria prevalence from the spatial malaria data using the ESRI ArcGIS© software and the process of merging them with each country's household data set. This process served as a type of introduction of the reader to the use of the ESRI ArcGIS© software. We also carried out covariate

selection algorithm on our chosen variables, after which the Markov chain Monte Carlo diagnostics were explored. We recapitulate our results by connecting them to our research objectives and questions thus:

Objective 1: *To estimate the technical efficiency of farmers due to malaria in our study areas*

The results of the composed-error model shows that farmers in Nigeria have a mean efficiency value of 27%, 89% in Ethiopia, and, 86% in Tanzania. From our result, the Ethiopian farmers are the most efficient of all farmers in the three countries. Linking the first objective to our first research question which is: "*Does malaria impact significantly on the efficiency of the farmers in each of the study areas*"?

Our results show that malaria does not significantly impact on the efficiency of the farmers in Nigeria, and Tanzania, but it impacts significantly on the efficiency of farmers in Ethiopia. The insignificance of malaria in Nigeria and Tanzania may due to the fact the farmers may have developed a certain level of tolerance to the disease. However, it is worthy of note that apart from Nigeria, the remaining two countries have shown high efficiency values, this could as a result of the government of these two countries have tried to reduce the incidence of the disease in their areas. This is evident from our chapter two, which states that, Ethiopia and Tanzania are one of the two countries in Africa have recorded successes in the area of eradication of the disease.

Objective 2: *To estimate how much the households are willing to pay for malaria abatement in our study areas*

The results of our analyses showthat on the average households are willing to pay in one farming year about *US$*0.1 for malaria abatement in Nigeria, *US$*0.98 in Ethiopia, and, *US$*0.07 in Tanzania. We also reported the individual country's posterior distribution.

The provision of the posterior distribution answers the second research question which is: "*How much are the households willing to pay for malaria abatement in our study areas*". The result of the posterior distribution of the three countries in our study shows that farmers in Ethiopia are willing to pay more for malaria abatement than the other two countries in our study. A plausible answer to this is that farmers in Ethiopia operate on the smallest piece of land out of farmers in the three countries we surveyed. Also, they operate on the least fertile lands out of the three countries. Thus, it is expected that they will be willing to invest in any programme that will increase their productivity and efficiency.

We also expect that Nigerian farmers should be willing to pay more for malaria prevention because it has the highest incidence of the disease in Africa, but this is not the case, this may be due to measurement error with regards to the malaria variable. This is because the malaria data did not have a one to one connection with individual household in our datasets. The malaria data used was based on the measurement of the risk that the farmer will be affected. Another plausible reason may be because the disease does not affect their productivity, although this needs to be investigated further.

In Tanzania, the malaria variable has the expected sign, but was not significant. The non-significance might as well be due to measurement errors which we have asserted in the last paragraph.

Objective 3: To make predictive inference on the future productivity of the due to malaria

Our results provide the mean predictive value for the composed error model we analysed for each country. This value is multiplied against the prevailing price per unit of agricultural staple and the amount that households are willing to pay will be provided in each of the countries. This also answers the last research question which is: "*Can we place a value on the amount the household is willing to pay in the future*" and our answer is "yes", we can place a value on the amount that the household is willing to pay in the future.

Overall, our method may not have followed the expected results, especially in Nigeria but we believe that following all the Bayesian steps we have stated a well behaved data should yield reliable results. Considering these limitations, we would say, we have been able to answer our research questions and have also been able to show that a method of measuring household willingness to pay is possible using a market based approach.

We have attempted to summarise our results in this section. The next section aims to provide the policy implication of this research.

10.2. Policy Implications

This study provides insight for policy making where the willingness-to- pay values can be used by decision makers in setting minimum prices for whatever prophylactic measures they want to introduce into the market. This point is supported by the fact that a lot of the successes achieved in reducing the gap between the *ownership* and *use* of some of the prophylactic measures introduced by policy makers are as a result of individuals purchasing these measures from retail shops instead of its distribution free of charge (see WHO 2012 document p. 12 for further exposition). Thus, our result will help to improve the use-efficiency of prophylactic measures.

Other things being equal, the resultant effect of the increase in use-efficiency of the prophylactic drugs is the reduction in the number of absentee farmers on the farms in our sample areas. This in-turn should lead to improved productivity and efficiency (especially in countries where malaria significantly affects efficiency) of the farming households. Increased productivity is expected to result in increased Gross Domestic Product; increased per capita income, expand the size of land cultivated, and, increase the standard of living of farmers through improved health. It is also expected to cause the farmer to switch from less profitable and less labour intensive crops to more profitable and more labour intensive crops.

The provision of a willingness-to-pay distribution in our results can help policy makers introduce gradual repayment plans for farmers that can not afford to pay for the preventive measure. The repayment plan should be devised in such a way that it suits the

farmers' income and his ability to pay for the preventive measure. The introduction of the repayment scheme will also help to increase the distribution of whatever prophylactic measures policy makers have put in place.

Also, the introduction of the minimum pricing and gradual repayment scheme are expected to bring more private organisations into the production and distribution of these preventive measures. Other things being equals, this will improve the development and introduction of new anti-malaria products into the market, and, the expansion of the distribution capacity for prophylactic measures. Other things being equal, the resultant effect is the introduction of competition into the market for the production of prophylactic measures which is expected to result in a price drop for anti-malaria measures. As a result, competing brands will become accessible and affordable to the farmers, which will lead to increased sales for the manufacturers. Also, this measure is expected to reduce "rent-seeking" behaviour among the manufacturers of these prophylactic measures.

The introduction of minimum prices is also expected to result into the equitable distribution of the prophylactic measures, this is because, if every farmer pays a certain minimum price for a prophylactic measure, then, it is likely to reduce the problem of a particular region or local government being favoured during the distribution of the measures. This is also expected to prevent certain groups of farmers from enjoying access to the preventive measures.

We have presented the implication of our results on ploicy making, in the next section, we highlight the merits of this research.

10.3. Merits of the Research

The primary merit of this research lies in the introduction and illustration of a market-based approach at arriving at a reliable amount farmers are willing to pay for malaria abatement. The provision of a reliable willingness-to-pay estimate has been a major course for concern among researchers in the literature. This is because of the numerous problems associated with the use of the non-market based method of stated preference measure in arriving at a measure of Willingness-To-Pay (Bryan and Dolan 2004 lists these problems). Also, added to the problems associated with the stated preference approach, is the fact that the amount arrived at by the method is greatly subjective, however, our method is objective. Our method can also help to determine the total amount the farmer is willing to pay over his lifetime, if his total life span is given.

Our method is also a novel idea in the area of economics of disease epidemiology by successfully introducing a method of arriving at this measure for a complicated disease like malaria. Thus, our method will go a long way in helping to eradicate malaria and similar diseases from the earth's surface. Our method will also aid the attainment of one of the new sustainable development goals of the United Nations, which are to end hunger and promote sustainable agriculture, ensuring healthy lives and promoting well-being for

all at all ages, as well as, ensuring sustainable consumption and production pattern by the year 2030.

Our method also serves as one of the few literature that has successfully combined and utilized the household and composed error models. The combination of these models, as well as, the utilization of the *envelope theorem*, we believe, will be useful in both theoretical and empirical literature of the future.

Our research has added to the little body of literature that has successfully combined spatial and survey data sets in a particular research. Our use of the spatial malaria data also helps to introduce the use of the ESRI ArcGIS© software to prospective users of the software.

Apart from Holloway et al. (2005), our study adds to to the few studies in Bayesian Econometrics literature that successfully renders a pedagogic presentation of the exponential composed-error model as well as the successful use of the Bayesian method of covariate selection. Our study further adds to argument on which part certain variables go in the composed error model.

The quantitative application of the more accurate malaria prevalence data in-lieu of the popular hospital-based malaria incidence data is also major contribution of this research to the literature on health economics.

The merits of this research have been stated in this section. In the next section, we emphasize the limitations and suggestion for future research.

10.4. Limitations and Suggestions for Future Research

Our research is not devoid of any limitations. The mitigation of these limitations will result in more reliable results.

A major limitation of this research is the fact that the malaria variable used did not directly relate to individual households in the countries investigated. This is possibly the reason why some of our results were not in line with expectations. As a result, we suggest that in the future, the malaria variable collected should be collected per household in the sample and not just the malaria prevalence value of the area where the households are located.

Also, we did not have enough time to investigate the level of spatial interaction, if any, that exists between households in a particular country and between the countries in our study. This is expected to throw more light on how much this affects the willingness-to-pay value.

As pointed out in Chapter 7, the lack of panel data in two of the countries we investigated precludes proper comparison of the willingness-to-pay values between these countries. Thus, a repeat of this study when panel data is available in Nigeria and Tanzania is there-

fore advised. The lack of collection of the same variables from the three different countries also precludes accurate comparison of our result among the three different countries.

A replication of this study in regions of unstable malaria will help policy makers understand how much households are willing to pay in case of an out-break of the disease in these regions. Also, as it was observed in our conceptual framework, the precision of our calculation is highly dependent on the accuracy and 'cleanliness' of the data, we suggest that 'cleaner' and more accurate data should be employed in the future.

Finally, It will also be of interest to seek the application of our method in other areas of economics like environmental and resource economics.

Bibliography

Abdulai, A. and Eberlin, R.: 2001, Technical efficiency during economic reform in nicaragua: evidence from farm household survey data, *Economic Systems* **25**(2), 113–125.

URL: *http://www.sciencedirect.com/science/article/pii/S0939362501000103*

Abdulai, A. and Huffman, W.: 2000, Structural adjustment and economic efficiency of rice farmers in northern ghana, *Economic Development and Cultural Change* **48**(3), 503–520.

URL: *http://www.jstor.org/stable/10.1086/452608*

Acemoglu, D., Johnson, S. and Rosinson, J. A.: 2001, The colonial origins of comparative development: An empirical investigation., *American Economic Review* **91**(5), 1369.

Adams, J.: 1986, Peasant rationality: Individuals, groups, cultures, *World Development* **14**(2), 273–282.

URL: *http://www.sciencedirect.com/science/article/pii/0305750X86900598*

Adegeye, A. J. and Ditto, J. S.: 1985, *Essentials of Agricultural Economics*, 2nd edn, lmpact Publishers lbadan Nigeria.

Ahmad, M. and Bravo-Ureta, B. E.: 1996, Technical efficiency measures for dairy farms using panel data: A comparison of alternative model specifications, *Journal of Productivity Analysis* **7**(4), 399–415.

Ahmad, R.: 1976, On the multivariatek-sample problem and the generalization of the

kolmogorov-smirnov-test, *Annals of the Institute of Statistical Mathematics* **28**(1), 259–265.

Aigner, D. J. and Chu, S. F.: 1968, On estimating the industry production function, *The American Economic Review* **58**(4), 826–839.
URL: *http://www.jstor.org/stable/1815535*

Aigner, D., Lovell, C. and Schmidt, P.: 1977, Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics* **6**, 21–37.

Ajani, O. and Ashagidigbi, W.: 2008, Effect of malaria on rural households' farm income in oyo state, nigeria, *African Journal of Biomedical Research* **11**.

Ajani, O. and Ugwu, P.: 2008, Impact of adverse health on agricultural productivity of farmers in kainji basin north-central nigeria using a stochastic production frontier approach, *Trends In Agricultural Economics* **1**(1), 1–7.

Ajetunmobi, W., Orimadegun, A., Brown, B., Afolabi, N., Olabiyi, F., Anetor, J., Omokhodion, S., Osinusi, K. and Others: 2012, Haeglobunuria amomg children with severe malaria attending tertiary care in ibadan, nigeria, *Malaria Journal* **11**(336).

Ajibefun, I. A.: 2002, Analysis of policy issues in technical efficiency of small scale farmers using the stochastic frontier production function: With application to nigerian farmers.
URL: *http://ageconsearch.umn.edu/bitstream/7015/2/cp02aj01.pdf*

Ajibefun, I. A. and Abdulkadri, A. O.: 1999, An investigation of technical inefficiency of production of farmers under the national directorate of employment in ondo state, nigeria, *Applied Economics Letters* **6**(2), 111–114. doi:.

Alaba, O. A. and Alaba, O. B.: 2009, Malaria in rural nigeria: Implications for the millennium development goals, *African Development Review* **21**(1), 73–85.

Alene, A. D. and Hassan, R. M.: 2006a, Erratum: The efficiency of traditional and hybrid maize production in eastern ethiopia: An extended efficiency decomposition approach, *Journal of African Economies* **15**(2), i–xxvii.

Alvarez, A., Perez, y. L. and Schmidt, P.: 2003, The relative importance of luck and technical efficiency in a fishery.

  URL: *http://uniovi.academia.edu/leviperez/Papers, www.uniovi.es/economia/edp.htm*

Aly, H., Belbase, K., Grabowski, R. and Krafrt, S.: 1987, The technical efficiency of illinois grain farms: An application of a ray-homothetic production function, *Southern Journal of Agricutural Economics* pp. 69 – 78.

  URL: *http://kolakaj.persiangig.com/ppt2.pdf*

Amacher, G. S., Hyde, W. F. and Kanel, K. R.: 1996, Household fuelwood demand and supply in nepal's tarai and mid-hills: Choice between cash outlays and labor opportunity, *World Development* **24**(11), 1725–1736.

Amara, N., Traorã, N., Landry, R. and Remain, R.: 1999, Technical efficiency and farmers' attitudes toward technological innovation: The case of the potato farmers in quebec, *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* **47**(1), 31–43.

  URL: *http://dx.doi.org/10.1111/j.1744-7976.1999.tb00214.x*

Amaza, P. S. and Olayemi, J. K.: 2002, Analysis of technical inefficiency in food crop production in gombe state, nigeria, *Applied Economics Letters* **9**(1), 51–54. doi:.

Amemiya, T.: 1974, Multivariate regression and simultaneous equation models when the dependent variables are truncated normal, *Econometrica: Journal of the Econometric Society* pp. 999–1012.

Amos, T.: 2007, An analysis of productivity and technical efficiency of smallholder cocoa farmers in nigeria, *Journal of Social Sciences* **15**(2), 127–133.

Apps, P.: 1982, Institutional inequality and tax incidence, *Journal of Public Economics* **18**(2), 217–242.

Apps, P. F. and Rees, R.: 1988, Taxation and the household, *Journal of Public Economics* **35**(3), 355 – 369.

  URL: *http://www.sciencedirect.com/science/article/pii/0047272788900370*

Apps, P. and Jones, G.: 1986, Selective taxation of couples, *Journal of Economics* **5**(1), 1–15.

Apps, P. and Rees, R.: 1996, Labour supply, household production and intra-family welfare distribution, *Journal of Public Economics* **60**(2), 199–219.

Areal, F. J., Tiffin, R. and Balcombe, K.: 2012, Farm technical efficiency under a tradable milk quota system, *Journal of Dairy Science* **95**(1), 50–62.

Ashenfelter, O. and Heckman, J.: 1974, The estimation of income and substitution effects in a model of family labor supply, *Econometrica* **42**(1), pp. 73–85.
   URL: *http://www-jstor-org.idpproxy.reading.ac.uk/stable/1913686*

Atkinson, S. E. and Cornwell, C.: 1993, Measuring technical efficiency with panel data: a dual approach, *Journal of Econometrics* **59**(3), 257–261.

Atkinson, S. E. and Cornwell, C.: 1994, Parametric estimation of technical and allocative inefficiency with panel data, *International Economic Review* pp. 231–243.

Audibert, M., Mathonnat, J. and Henry, M.-C.: 2003, Social and health determinants of the efficiency of cotton farmers in northern cote d'ivoire, *Social Science and amp Medicine* **56**(8), 1705–1717.
   URL: *http://www.sciencedirect.com/science/article/pii/S0277953602001648*

Azzi, C. and Ehrenberg, R.: 1975, Household allocation of time and church attendance, *The Journal of Political Economy* pp. 27–56.

Badiane, O. and Ulimwengu, J.: 2013, Malaria incidence and agricultural efficiency in uganda, *Agricultural Economics* **44**(1), 15–23.

Balcombe, K., Fraser, I. and Kim, J. H.: 2006, Estimating technical efficiency of australian dairy farms using alternative frontier methodologies, *Applied Economics* **38**(19), 2221–2236.

Ball, R. and Pounder, L.: 1996, "efficient but poor" revisited, *Economic Development and Cultural Change* **44**(4), pp.735–760.
   URL: *http://www.jstor.org/stable/1154350*

Banker, R. D., Charnes, A. and Cooper, W. W.: 1984, Some models for estimating technical AND SCale inefficiencies in data envelopment analysis., *Management Science* **30**(9), 1078–1092.

Barnett, W. A.: 1977, Pollak and wachter on the household production function approach, *Journal of Political Economy* **85**(5), 1073–1082.
  URL: *http://www.jstor.org/stable/1830347*

Barnum, H. and Squire, L.: 1979a, An econometric application of the theory of the farm-household, *Journal of Development Economics* **6**, 79–102.

Barnum, H. and Squire, L.: 1979b, *A Model of an Agricultural Household. Theory and Evidence*, Vol. 27 of *World Bank Staff Occassional Papers (IBRD)*, The World Bank.

Battese, G. E. and Coelli, T. J.: 1988, Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data, *Journal of Econometrics* **38**(3), 387–399.
  URL: *http://www.sciencedirect.com/science/article/pii/030440768890053X*

Battese, G. E. and Coelli, T. J.: 1992, Frontier production functions, technical efficiency and panel data: With application to paddy farmers in india, *Journal of Productivity Analysis* **3**(1), 153–169.

Battese, G. E. and Coelli, T. J.: 1995, A model for technical inefficiency effects in a stochastic frontier production function for panel data, *Empirical Economics* **20**(2), 325–332.

Battese, G. E., Coelli, T. J. and Colby, T.: 1989, *Estimation of frontier production functions and the efficiencies of Indian farms using panel data from ICRISAT's village level studies*, Department of Econometrics, University of New England.

Battese, G. E. and Corra, G. S.: 1977, Estimation of a production frontier model: With application to the pastoral zone of eastern australia, *Australian Journal of Agricultural and Resource Economics* **21**(3), 169–179.

Bauer, P. W.: 1990, Recent developments in the econometric estimation of frontiers, *Journal of econometrics* **46**(1), 39–56.

Bauer, P. W., Berger, A. N. and Humphrey, D. B.: 1993, Efficiency and productivity growth in us banking, *The measurement of productive efficiency: Techniques and applications* pp. 386–413.

Bauer, P. W. and Hancock, D.: 1993, The efficiency of the federal reserve in providing check processing services, *Journal of Banking & Finance* **17**(2 - 3), 287 – 311.
  URL: *http://www.sciencedirect.com/science/article/pii/037842669390033A*

Becker, G. S.: 1965, A theory of the allocation of time, *The Economic Journal* **75**(299), 493–517.
  URL: *http://www.jstor.org/stable/2228949*

Belbase, K. and Grabowski, R.: 1985, Technical efficiency in nepalese agriculture, *The Journal of Developing Areas* **19**(4), 515–526.
  URL: *http://www.jstor.org/stable/4191394*

Benjamin, D.: 1992, Household composition, labor markets, and labor demand: Testing for separation in agricultural household models, *Econometrica* **60**(2), 287–322.

Bera, A. K. and Sharma, S. C.: 1999, Estimating production uncertainty in stochastic frontier production function models, *Journal of Productivity Analysis* **12**(3), 187–210.

Best, N., Cowles, M. K. and Vines, K.: 1995, Coda* convergence diagnosis and output analysis software for gibbs sampling output version 0.30, *MRC Biostatistics Unit, Cambridge* **52**.

Bezemer, D., Balcombe, K., Davis, J. and Fraser, I.: 2005, Livelihoods and farm efficiency in rural georgia, *Applied Economics* **37**(15), 1737–1745. doi:.

Billingsley, P.: 1971, *Weak Convergence of Measures.*

Binam, J. N., Tonyã, J., Wandji, N., Nyambi, G. and Akoa, M.: 2004, Factors affecting the technical efficiency among smallholder farmers in the slash and burn agriculture zone of cameroon, *Food Policy* **29**(5), 531–545.

Bjurek, H., Hjalmarsson, L. and Forsund, F. R.: 1990, Deterministic parametric and nonparametric estimation of efficiency in service production: A comparison, *Journal of Econometrics* **46**(1 and 2), 213–227.

Blundell, R. W. and Walker, I. F.: 1981, *The Demand for Goods and Leisure with Female Labour Supply and Rationing: Testing for Separability*, Department of Economics, University of British Columbia.

Boone, E. L., Merrick, J. R. and Krachey, M. J.: 2014, A hellinger distance approach to mcmc diagnostics, *Journal of Statistical Computation and Simulation* **84**(4), 833–849.

Breman, J., Egan, A. and Keusch, G.: 2001, The intolerable burden of malaria: A new look at the numbers, *American Journal of Tropical Medicine and Hygiene* **64**(Supplement,1), 4 – 7i.

Breman, J. G., Alilio, M. S., Mills, A., Breman, J. G. and Alilio, M. S.: 2004, Conquering the intolerable burden of malaria: What is new, what is needed: A summary.

Brãnnlund, R. and Nordstrãm, J.: 2004, Carbon tax simulations using a household demand model, *European Economic Review* **48**(1), 211–233.
  URL: *http://www.sciencedirect.com/science/article/pii/S0014292102002635*

Brooks, S., Dellaportas, P. and Roberts, G.: 1997, A total variation method for diagnosing convergence of mcmc algorithms, *Journal of Computational and Graphical Statistics* **6**, 251–265.

Brooks, S., Giudici, P. and Philippe, A.: 2003, Nonparametric convergence assessment for mcmc model selection, *Journal of Computational and Graphical Statistics* **12**(1), 1–22.

Brooks, S. P. and Gelman, A.: 1998, General methods for monitoring convergence of iterative simulations, *Journal of Computational and Graphical Statistics* **7**(4), 434–455.

Brooks, S. P. and Giudici, P.: 2000, Markov chain monte carlo convergence assessment via two-way analysis of variance, *Journal of Computational and Graphical Statistics* **9**(2), 266–285.

Brooks, S. P. and Roberts, G. O.: 1998, Convergence assessment techniques for markov chain monte carlo, *Statistics and Computing* **8**(4), 319–335.

Brown, B. and LeVasseur, M.: 2006, National geographic: Physical geography, National Geographic Society.
URL: *http://www.nationalgeographic.com/xpeditions/guides/physicalafricaguide.pdf*

Brown, M. and Manser, M.: 1978, Neoclassical and bargaining approaches to household decision-making-with application to the household labor supply decision, *Technical report*, Discussion Paper.

Browning, M., Deaton, A. and Irish, M.: 1985, A profitable approach to labor supply and commodity demands over the life-cycle, *Econometrica: Journal of the Econometric Society* pp. 503–543.

Browning, M. and Meghir, C.: 1991, The effects of male and female labor supply on commodity demands, *Econometrica: Journal of the Econometric Society* pp. 925–951.

Bryan, S. and Dolan, P.: 2004, Discrete choice experiments in health economics, *The European Journal of Health Economics* **5**.

Byrnes, P., Färe, R. and Grosskopf, S.: 1984, Measuring productive efficiency: An application to illinois strip mines, *Management Science* **30**(6), 671–681.

Casella, G. and George, E. I.: 1992, Explaining the gibbs sampler, *The American Statistician* **46**(3), 167–174.
URL: *http://www.jstor.org/stable/2685208*

Chambers, R. G.: 1988, *Applied Production Analysis*, Cambridge University Press.

Charnes, A., Cooper, W. W. and Rhodes, E.: 1978, Measuring the efficiency of decision making units, *European Journal of Operational Research* **2**(6), 429–444.
URL: *http://www.sciencedirect.com/science/article/pii/0377221778901388*

Chavas, J.-P.: 2006, Dynamic considerations in nutrition and food demand.

Chen, L., Heerink, N. and van den Berg, M.: 2006, Energy consumption in rural china: A

household model for three villages in jiangxi province, *Ecological Economics* **58**(2), 407–420.

Chen, Z., Huffman, W. E. and Rozelle, S.: 2009, Farm technology and technical efficiency: Evidence from four regions in china, *China Economic Review* **20**(2), 153–161.

Chennareddy, V.: 1967, Production efficiency in south indian agriculture, *Journal of Farm Economics* **49**(4), 816–820.
  URL: *http://www.jstor.org/stable/1236938*

Chiang, A. and Wainwright, K.: 2005, *Fundamental of Mathematical Economics*, 4th edn, McGraw-Hill/Irwin.

Chiappori, P.-A.: 1988, Nash-bargained households decisions: A comment, *International Economic Review* **29**(4), pp.791–796.
  URL: *http://www-jstor-org.idpproxy.reading.ac.uk/stable/2526833*

Chib, S.: 1992, Bayes inference in the tobit censored regression model, *Journal of Econometrics* **51**(1), 79–99.

Chib, S.: 1995, Marginal likelihood from the gibbs output, *Journal of the American Statistical Association* **90**(432), 1313–1321.

Chib, S. and Greenberg, E.: 1995, Understanding the metropolis-hastings algorithm, *The American Statistician* **49**(4), 327–335.
  URL: *http://www.jstor.org/stable/2684568*

Chib, S. and Jeliazkov, I.: 2001, Marginal likelihood from the metropolis-hastings output, *Journal of the American Statistical Association* **96**(453), 270–281.
  URL: *http://www.jstor.org/stable/2670365*

Chima, R., Goodman, A. and Mills, A.: 2003, The economic impact of malaria in africa: A critical review of evidence, *Health Policy* **63**, 17 – 36.

Chirwa, E.: 2003, Sources of technical efficiency among smallholder maize farmers in southern malawi, *Technical Report Wadonda Consult Working Paper WC/01/03*, University of Malawi.

Christensen, L. R. and Greene, W. H.: 1976, Economies of scale in u.s. electric power generation, *Journal of Political Economy* **84**(4), 655–676.

URL: *http://www.jstor.org/stable/1831326*

Chuma, J., Okungu, V. and Molyneux, C.: 2010, The ecoomic costs of malaria in four kenyan districts: do household costs differ by disease endemicity, *Malaria Journal* **9**(149).

URL: *www.malariajournal.com/content/9/1/149*

Coelli, T. J.: 1995, Recent development in frontier modelling and efficiency measurement, *The Australian Journal of Agricultural and Resource Economics* **39**(3), 219–245.

URL: *http://ageconsearch.umn.edu/bitstream/22681/1/39030219.pdf*

Coelli, T. J., Prasada Rao, D., O'Donnell, C. and Battese, G. E.: 2005, *An Introduction to Efficiency and Productivity Analysis*, 2nd edn, Springer.

Coelli, T., Rahman, S. and Thirtle, C.: 2003, A stochastic frontier approach to total factor productivity measurement in bangladesh crop agriculture, 1961- 92, *Journal of International Development* **15**(3), 321–333.

Comanor, W. S. and Leibenstein, H.: 1969, Allocative efficiency, x-efficiency and the measurement of welfare losses, *Economica* **36**(143), 304–309.

Conover, W. J.: 1965, Several k-sample kolmogorov-smirnov tests, *The Annals of Mathematical Statistics* **36**(3), 1019–1026.

Conover, W. J.: 1967, A k-sample extension of the one-sided two-sample smirnov test statistic, *The Annals of Mathematical Statistics* **38**(6), 1726–1730.

Cook, S.: 1966, The obsolete "anti-market" mentality: A critique of the substantive approach to economic anthropology, *American Anthropologist* **68**(2), 323–345.

Cornes, R.: 1992, *Duality and Modern Economics*, Cambridge University Press.

Cornia, G. A.: 1985, Farm size, land yields and the agricultural production function: An analysis for fifteen developing countries, *World Development* **13**, 513–534.

Cornwell, C. and Schmidt, P.: 2008, Stochastic frontier analysis and efficiency estimation, *in* l. Mãtyãs and P. Sevestre (eds), *The Econometrics of Panel Data*, Vol. 46 of *Advanced Studies in Theoretical and Applied Econometrics*, Springer Berlin Heidelberg, pp. 697–726.

Cornwell, C., Schmidt, P. and Sickles, R. C.: 1990, Production frontiers with cross-sectional and time-series variation in efficiency levels, *Journal of Econometrics* **46**(1-2), 185–200.
URL: *http://www.sciencedirect.com/science/article/pii/030440769090054W*

Cowles, M. K. and Carlin, B. P.: 1996, Markov chain monte carlo convergence diagnostics: A comparative review, *Journal of the American Statistical Association* **91**(434), 883–904.
URL: *http://www.jstor.org/stable/2291683*

Croppenstedt, A. and Muller, C.: 2000, The impact of farmers' health and nutritional status on their productivity and efficiency: Evidence from ethiopia, *Economic Development and Cultural Change* **48**(3), 475–502.
URL: *http://www.jstor.org/stable/10.1086/452607*

Dalton, G.: 1961, Economic theory and primitive society, *American Anthropologist* **63**(1), 1–25.

Dalton, T. J.: 2004, A household hedonic model of rice traits: economic values from farmers in west africa, *Agricultural Economics* **31**(2-3), 149–159.

Darling, D. A.: 1957, The kolmogorov-smirnov, cramer-von mises tests, *The Annals of Mathematical Statistics* **28**(4), 823–838.
URL: *http://www.jstor.org/stable/2237048*

Daunfeldt, S.-O. and Hellström, J.: 2007, Intra-household allocation of time to household production activities: Evidence from swedish household data, *Labour* **21**(2), 189–207.

Dawson, P. J., Lingard, J. and Woodford, C. H.: 1991, A generalized measure of farm-specific technical efficiency, *American Journal of Agricultural Economics* **73**(4), 1098–

1104.

   URL: *http://ajae.oxfordjournals.org/content/73/4/1098.abstract*

De Janvry, A., Fafchamps, M. and Sadoulet, E.: 1991, Peasant household behaviour with missing markets: some paradoxes explained, *The Economic Journal* **101**(409), 1400–1417.

De Janvry, A. and Sadoulet, E.: 1995, Household modeling for the design of poverty alleviation strategies, *Revue Economie du Developpement* **3**, 3–23.

De Janvry, A., Sadoulet, E., Fafchamps, M. and Raki, M.: 1992, Structural adjustment and the peasantry in morocco: A computable household model, *European Review of Agricultural Economics* **19**(4), 427–453.

   URL: *http://erae.oxfordjournals.org.idpproxy.reading.ac.uk/content/19/4/427.abstract*

Dean, E. R.: 1965, Economic analysis and african responses to price, *Journal of Farm Economics* **47**(2), pp.402–409.

   URL: *http://www.jstor.org/stable/1236585*

Deaton, A. and Muellbauer, J.: 1980, *Economics and Consumer Behavior*, Cambridge University Press.

Debreu, G.: 1951, The coefficient of resource utilization, *Econometrica* **19**(3), pp. 273–292.

   URL: *http://www.jstor.org/stable/1906814*

Del Moral, P., Doucet, A. and Jasra, A.: 2006, Sequential monte carlo samplers, *Journal of the Royal Statistical Society* **68**(3), 411–436.

Dercon, S. and Hoddinott, J.: 2011, The ethiopian rural household surveys 1989 - 2009, *Technical report*, Addis Ababa University(Economics Department), Centre for Study of African Economics (OXford) and International Food Policy Research Institute.

Didelot, X., Everitt, R. G., Johansen, A. M. and Lawson, D. J.: 2011, Likelihood-free estimation of model evidence, *Bayesian Anal.* **6**(1), 49–76.

Doiron, D. and Kalb, G.: 2004, Demands for childcare and household labour supply in australia.

Edmeades, S. O.: 2003, Variety choice and attribute trade-offs within the framework of agricultural household models: The case of bananas in uganda.

Edwards, L. N. and Grossman, M.: 1979, Adolescent health, family background, and preventive medical care.

Elhorst, J. P.: 1994, Firm-household interrelationships on dutch dairy farms, *European Review of Agricultural Economics* **21**(2), 259–276.
  URL: *http://erae.oxfordjournals.org.idpproxy.reading.ac.uk/content/21/2/259.abstract*

Ethiopia Ministry of Health: 2006, National 5-year strategic plan for malaria prevention and control in ethiopia 2006 - 2010, *Technical report*, Ministry of Health, Addis Ababa.

Ettling, M., McFarland, D., Schultz, L. and Chitsulo, L.: 1994, Economic impact of malaria in malawian households, *Tropical Medicine and Parasitology* **45**(1), 74–79.
  URL: *http://www.ncbi.nlm.nih.gov/pubmed/8066390*

Fafchamps, M.: 1993, Sequential labor decisions under uncertainty: An estimable household model of west-african farmers, *Econometrica* **61**(5), pp.1173–1197.
  URL: *http://www.jstor.org/stable/2951497*

Fare, R., Grosskopf, S., Logan, J. and Lovell, C. K.: 1985, Measuring efficiency in production: With an application to electric utilities, *Managerial Issues in Productivity Analysis*, Springer, pp. 185–214.

Fare, R., Grosskopf, S. and Lovell, C. K.: 1994, *Production frontiers*, Cambridge University Press.

Farrell, M. J.: 1957, The measurement of productive efficiency, *Journal of the Royal Statistical Society. Series A (General)* **120**(3), pp. 253–290.
  URL: *http://www.jstor.org/stable/2343100*

Fernändez, C., Koop, G. and Steel, M.: 2000, A Bayesian analysis of multiple-output production frontiers, *Journal of Econometrics* **98**(1), 47–79.

Fernändez, C., Koop, G. and Steel, M. F. J.: 2002, Multiple-output production with undesirable outputs, *Journal of the American Statistical Association* **97**(458), 432–442. doi:.

Fernändez, C., Koop, G. and Steel, M. F. J.: 2005, Alternative efficiency measures for multiple-output production, *Journal of Econometrics* **126**(2), 411–444.

Fernandez, C., Osiewalski, J. and Steel, M. J.: 1997, On the use of panel data in stochastic frontier models with improper priors, *Journal of Econometrics* **79**(1), 169–193.
    URL: *http://www.sciencedirect.com/science/article/pii/S0304407697880505*

Flegal, J. M., Haran, M. and Jones, G. L.: 2008, Markov chain monte carlo: Can we trust the third significant figure?, *Statist. Sci.* **23**(2), 250–260.
    URL: *http://dx.doi.org/10.1214/08-STS257*

Flores-Lagunes, A., Horrace, W. C. and Schnier, K. E.: 2007, Identifying technically efficient fishing vessels: a non-empty, minimal subset approach, *Journal of Applied Econometrics* **22**(4), 729–745.

Fãre, R. and Lovell, C. K.: 1978, Measuring the technical efficiency of production, *Journal of Economic Theory* **19**(1), 150–162.

Fried, H. and Lovell, C.: 2008, The measurement of productive efficiency and productivity change.

Fãrsund, F. R. and Hjalmarsson, L.: 1979, Generalised farrell measures of efficiency: An application to milk processing in swedish dairy plants, *The Economic Journal* **89**(354), 294–315.
    URL: *http://www.jstor.org/stable/2231603*

Furtan, W. H., van Kooten, G. C. and Thompson, S. J.: 1985, The estimation of off-farm labour supply functions in saskatchewan, *Journal of Agricultural Economics* **36**(2), 211–220.
    URL: *http://dx.doi.org/10.1111/j.1477-9552.1985.tb00167.x*

Gallup, J. and Sachs, J.: 2001, The economic burden of malaria, *American Journal of Tropical Medicine and Hygiene* **64**(1), 85–96.

Garcia, I. and Marcuello, C.: 2001, A household model of charitable contributions and tax incentives, *Annals of Public and Cooperative Economics* **72**(2), 159–181.

Garren, S. T. and Smith, R. L.: 2000, Estimating the second largest eigenvalue of a markov transition matrix, *Bernoulli* **6**(2), 215–242.

Garśia, A. M.: 1970, *Topics in almost everywhere convergence*, Vol. 4, Rand McNally.

Gathon, H.-J. and Perelman, S.: 1992, Measuring technical efficiency in european railways: A panel data approach, *Journal of Productivity Analysis* **3**(1/2), 135–151.

Gelfand, A. E. and Dey, D. K.: 1994, Bayesian model choice: Asymptotics and exact calculations, *Journal of the Royal Statistical Society. Series B (Methodological)* **56**(3), 501–514.
URL: *http://www.jstor.org/stable/2346123*

Gelfand, A. E. and Smith, A. F. M.: 1990, Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association* **85**(410), 398–409.
URL: *http://www.jstor.org/stable/2289776*

Gelman, A., Carlin, J. and Rubin, D.: 2003, *Bayesian Data Analysis*, (2nd edition) edn, Chapman & Hall/CRC, New york.

Gelman, A. and Rubin, D. B.: 1992, Inference from iterative simulation using multiple sequences, *Statistical Science* **7**(4), 457–472.
URL: *http://www.jstor.org/stable/2246093*

Geman, S. and Geman, D.: 1984, Stochastic relaxation, gibbs distributions, and the bayesian restoration of images, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **PAMI-6**(6), 721–741.

Gentle, J.: 2003, *Random Number Generation and Monte Carlo Methods*, Springer Science & Business Media B.V.

Gething, P. W., Patil, A. P., Smith, D. L., Guerra, C. A., Elyazar, I., , G. L., Tatem, A. J., Hay, S. I. et al.: 2011, A new malaria map: Plasmodium falciparum endemicity in 2010, *Malaria Journal* **10**(378), 1475–2875.

Geweke, J.: 1999, On markov chain monte carlo methods for nonlinear and non-gaussian state-space models, *Communications in statistics. Simulation and Computation* **28**(4), 867–894.

Geweke, J. et al.: 1991, *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, Vol. 196, Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA.

Giakoumatos, S., Vrontos, I., Politis, D. and Dellaportas, P.: 1999, A markov chain monte carlo convergence diagnostic using subsampling, *Journal of Computational and Graphical Statistics* **8**(3), 431–451.

Gibbs, A. L.: 2000, Bounding the convergence time of the gibbs sampler in bayesian image restoration, *Biometrika* **87**(4), 749–766.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J.: 1996, Introducing markov chain monte carlo, *Markov chain Monte Carlo in practice* **1**, 19.

Goldman, F. and Grossman, M.: 1978, The demand for pediatric care: An hedonic approach, *The Journal of Political Economy* pp. 259–280.

Gong, B.-H. and Sickles, R. C.: 1989, Finite sample evidence on the performance of stochastic frontier models using panel data, *Journal of Productivity Analysis* **1**(3), 229–261.

Gong, L.: 2015, *Fixed-Width Stopping Procedures for Markov Chain Monte Carlo*, PhD thesis, University Of California Riverside, Department Of Applied Statistics.

Greene, W.: 1993, *The Econometric Approach to Efficiency Measurement in "The Measurement of Productive of Efficiency"*, Oxford University Press, New York, chapter 2, pp. 68 – 119.

Greene, W. H.: 1980, Maximum likelihood estimation of econometric frontier functions, *Journal of Econometrics* **13**, 27–56.

Greene, W. H.: 1990, A gamma-distributed stochastic frontier model, *Journal of Econometrics* **46**(1 and 2), 141–163.

Greene, W. H.: 1997, Frontier production functions, *Handbook of applied econometrics* **2**, 81–166.

Greene, W. H.: 2008, *The Measurement of productive Efficiency and Productivity growth*, 1st edn, Oxford University Press, chapter 2, pp. 92 – 250.

Hajek, A.: 2012, *Interpretation of Probability*, The Stanford Encyclopedia of Philosophy, summer 2012 edn, Stanford.

Hansen, B. E.: 2005, Challenges for econometric model selection, *Econometric Theory* **21**(1), 60–68.

Harris, M. C. A. B. I. E.: 1989, *Technical efficiency in Australian manufacturing industries*, 1 edn, Bureau of Industry Economics.

Hastings, W. K.: 1970, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* **57**(1), 97–109.
URL: *http://www.jstor.org/stable/2334940*

Hausman, J. A. and Taylor, W. E.: 1981, Panel data and unobservable individual effects, *Econometrica* **49**(6), pp.1377–1398.
URL: *http://www.jstor.org/stable/1911406*

Hay, S. I., Guerra, C. A., Gething, P. W., Patil, A. P., Tatem, A. J., Noor, A. M., Kabaria, C. W., Manh, B. H., Elyazar, I. R., , S. et al.: 2009, A world malaria map: Plasmodium falciparum endemicity in 2007, *PLoS medicine* **6**(3), e1000048.

Hay, S. I., Guerra, C. A., Tatem, A. J., Atkinson, P. M. and Snow, R. W.: 2005, Tropical infectious diseases: Urbanization, malaria transmission and disease burden in africa, *Nature Review Microbiology* **3**(1).
URL: *http://dx.doi.org/10.1038/nrmicro1069*

Hay, S. I., Okiro, E. A., Gething, P. W., Patil, A. P., Tatem, A. J., Guerra, C. A. and Snow, R. W.: 2010, Estimating the global clinical burden of plasmodium falciparum malaria in 2007, *PLoS medicine* **7**(6), 1000290.

Healy, M.: 2012, Sub-saharan africa: Physical geography, Economics and Geography: Harper College.
URL: *http://www.harpercollege.edu/mhealy/g101ilec/ssa/afd/afphys/afphysfr.htm*

Heckman, J. J.: 1979, Sample selection bias as a specification error, *Econometrica* **47**(1), pp.153–161.
URL: *http://www-jstor-org.idpproxy.reading.ac.uk/stable/1912352*

Heidelberger, P. and Welch, P. D.: 1981a, Adaptive spectral methods for simulation output analysis, *IBM Journal of Research and Development* **25**(6), 860–876.

Heidelberger, P. and Welch, P. D.: 1981b, A spectral method for confidence interval generation and run length control in simulations, *Commun. ACM* **24**(4), 233–245.

Heidelberger, P. and Welch, P. D.: 1983, Simulation run length control in the presence of an initial transient., *Operations Research* **31**(6), 1109 – 1144.

Heltberg, R., Arndt, T. C. and Sekhar, N. U.: 2000, Fuelwood consumption and forest degradation: A household model for domestic energy substitution in rural india, *Land Economics* **76**(2), 213–232.
URL: *http://www.jstor.org/stable/3147225*

Heshmati, A., Kumbhakar, S. C. and Hjalmarsson, L.: 1995, Efficiency of the swedish pork industry: A farm level study using rotating panel data 1976 to 1988, *European Journal of Operational Research* **80**(3), 519–533.
URL: *http://www.sciencedirect.com/science/article/pii/037722179400135*

Hill, J., Lines, J. and Rowland, M.: 2006, Insecticide-treated nets, *in* H. M. David (ed.), *Advances in Parasitology*, Vol. Volume 61, Academic Press, pp. 77–128.
URL: *http://www.sciencedirect.com/science/article/pii/S0065308X05610032*

Hjalmarsson, L., Kumbhakar, S. and Heshmati, A.: 1996, Dea, dfa and sfa: A comparison, *Journal of Productivity Analysis* **7**(2-3), 303–327.

Holloway, G. and Ehui, K. S.: 2001, Demand, supply and willingness-to-pay for extension services in an emerging-market setting, *American Journal of Agricultural Economics* **83**(3), 764–768.
  URL: *http://www.jstor.org/stable/1245114*

Holloway, G. and Paris, Q.: 2002, Production efficiency in the von liebig model, *American Journal of Agricultural Economics* **84**(5), pp.1271–1278.
  URL: *http://www.jstor.org/stable/1245058*

Holloway, G. and Tomberlin, D.: 2007, Bayesian ranking and selection of fishing boat efficiencies, *Marine Resource Economics* **21**, 415–432.

Holloway, G., Tomberlin, D. and Irz, X.: 2005, Hierarchical analysis of production efficiency in a coastal trawl fishery, *in* R. Scarpa and A. Alberini (eds), *Applications of Simulation Methods in Environmental and Resource Economics*, Vol. 6 of *The Economics of Non-Market Goods and Resources*, Springer Netherlands, pp. 159–185.

Holtz-Eakin, D., Newey, W. and Rosen, H. S.: 1988, Estimating vector autoregressions with panel data, *Econometrica* **56**(6), 1371–1395.
  URL: *http://www.jstor.org/stable/1913103*

Horrace, W. C.: 2005, On ranking and selection from independent truncated normal distributions, *Journal of Econometrics* **126**(2), 335–354.

Horrace, W. C. and Schmidt, P.: 1996, Confidence statements for efficiency estimates from stochastic frontier models, *Journal of Productivity Analysis* **7**(2-3), 257–282.

Horrace, W. C. and Schmidt, P.: 2000, Multiple comparisons with the best, with economic applications.

Hsiao, C.: 2003, *Analysis of Panel Data*, 2nd edn, Cambridge University Press, UK.

Hsu, E.: 1999, *Probability Theory and Applications*, Vol. 6, American mathematical Society.

Huang, C. J.: 1984, Estimation of stochastic frontier production function and technical inefficiency via the em algorithm, *Southern Economic Journal* **50**(3), 847–856.
  URL: *http://www.jstor.org/stable/1057997*

Huffman, W. E.: 1976, The productive value of human time in u.s. agriculture, *American Journal of Agricultural Economics* **58**(4), pp. 672–683.
  URL: *http://www.jstor.org/stable/1238810*

Innes, R.: 1993, Two-season subsistence farming, urban food subsidies and optimal agricultural policy, *Oxford Economic Papers* **45**(4), pp. 668–690.
  URL: *http://www.jstor.org/stable/2663710*

Jacoby, H. G.: 1993, Shadow wages and peasant family labour supply: An econometric application to the peruvian sierra, *The Review of Economic Studies* **60**(4), pp.903–921.
  URL: *http://www-jstor-org.idpproxy.reading.ac.uk/stable/2298105*

Jara-rojas, R., Bravo-ureta, B. E., Moreira, V. and Dãaz, J.: 2012, Natural resource conservation and technical efficiency from small-scale farmers in central chile.

Jensen, U.: 2000, Is it efficient to analyse efficiency rankings?, *Empirical Economics* **25**(2), 189–208.

Jimoh, A., Sofola, O. and Okorosobo, T.: 2007, Quantifying the conomic burden of malaria in nigeria using the willingness to pay approach, *Malaria Journal* **5**(6).
  URL: *www.resource-allocation.com/content/5/1/6*

Johansson, P. and Brafinnafis, K.: 1998, A household model for work absence, *Applied Economics* **30**(11), 1493–1503.

Johansson, P. and Palme, M.: 1996, Do economic incentives affect work absence? empirical evidence using swedish micro data, *Journal of Public Economics* **59**(2), 195 – 218.
  URL: *http://www.sciencedirect.com/science/article/pii/0047272795015019*

Johnson, V. E.: 1996, Studying convergence of markov chain monte carlo algorithms using coupled sample paths, *Journal of the American Statistical Association* **91**(433), 154–166.

Jondrow, J., Knox Lovell, C. A., Materov, I. S. and Schmidt, P.: 1982, On the estimation of technical inefficiency in the stochastic frontier production function model, *Journal of Econometrics* **19**(2 and 3), 233–238.
  URL: *http://www.sciencedirect.com/science/article/pii/0304407682900045*

Jones, G. L., Haran, M., Caffo, B. S. and Neath, R.: 2006, Fixed-width output analysis for markov chain monte carlo, *Journal of the American Statistical Association* **101**(476), 1537–1547.

Jorgenson, D. W. and Lau, L.: 1969, An economic theory of agrucultural household behaviour, *Technical report*, Paper read at the fourth Far Eastern Meeting of the Econometric Society.

Kalaitzandonakes, N. and Dunn, E.: 1995, Technical efficiency, managerial ability and farmer education in guatemalan corn production: A latent variable analysis, *Agricutulral Resource Economic Review* **24**, 36 – 46.

Kalirajan, K. P. and Flinn, J.: 1983, The measurement of farm specific technical efficiency, *Pakistan Journal of Applied Economics* **2**(2), 167–180.

Kalirajan, K. P. and Obwona, M. B.: 1994, Frontier production function: The stochastic coefficients approach, *Oxford Bulletin of Economics and Statistics* **56**(1), 87–96.

Kalirajan, K. P. and Shand, R.: 1999, Frontier production functions and technical efficiency measures, *Economics Surveys* **13**(2), 149–172.

Keiser, J., Caldas De Castro, M., Maltese, F., Bos, R., Tanner, M., Singer, B. and Utzinger, J.: 2005, Effect of irrigation and large dams on the burden of malaria on a global and regional scale, *Am J Trop Med Hyg* **72**(4), 392–406.

Key, N., Sadoulet, E. and Janvry, A. D.: 2000, Transactions costs and agricultural household supply response, *American Journal of Agricultural Economics* **82**(2), 245–259.
  URL: *http://ajae.oxfordjournals.org/content/82/2/245.abstract*

Kim, Y. . and Schmidt, P. .: 2000, A review and empirical comparison of Bayesian and

classical approaches to inference on efficiency levels in stochastic frontier models with panel data, *Journal of Productivity Analysis* **14**(2), 91.

Knight, F.: 1933, *The Economic Organization*, 1 edn, The University of Chicago Press.

Koop, G.: 2001, Cross-sectoral patterns of efficiency and technical change in manufacturing, *International Economic Review* **42**(1), 73–103.

Koop, G.: 2002, Comparing the performance of baseball players: A multiple-output approach, *Journal of the American Statistical Association* **97**(459), pp.710–720.
URL: *http://www-jstor-org.idpproxy.reading.ac.uk/stable/3085707*

Koop, G.: 2003, *Bayesian Econometrics*, 1st edn, John Wiley & Sons Ltd, Chichester.
URL: *www.wiley.co.uk/koopbayesian*

Koop, G., Osiewalski, J. and Steel, M. F. J.: 1994, Bayesian efficiency analysis with a flexible form: The AIm cost function, *Journal of Business & Economic Statistics* **12**(3), 339–346.
URL: *http://www.jstor.org/stable/1392090*

Koop, G., Osiewalski, J. and Steel, M. F. J.: 1997, Bayesian efficiency analysis through individual effects: Hospital cost frontiers, *Journal of Econometrics* **76**(1- 2), 77–105.
URL: *http://www.sciencedirect.com/science/article/pii/0304407695017836*

Koop, G., Osiewalski, J. and Steel, M. F. J.: 1999, The components of output growth: A stochastic frontier analysis., *Oxford Bulletin of Economics & Statistics* **61**(4), 455.

Koop, G., Steel, M. and Osiewalski, J.: 1994, Posterior analysis of stochastic frontier models using gibbs sampling, *CORE Discussion Papers 1994061*, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).
URL: *http://EconPapers.repec.org/RePEc:cor:louvco:1994061*

Koopmans, T. C.: 1951, Analysis of production as an efficient combination of activities, *Activity analysis of production and allocation* **13**, 33–37.

Kopp, R. J.: 1981, The measurement of productive efficiency: A reconsideration, *The*

Quarterly Journal of Economics **96**(3), 477–503.

  URL: *http://qje.oxfordjournals.org/content/96/3/477.abstract*

Krishna, R.: 1964, Theory of the firm: Rapporteurs report, *Indian Economic Journal* **11**, 514–525.

Kumbhakar, S.: 1989, Modelling technical and allocative inefficiency in a translog production function, *Economic Letters* **31**, 119–123.

Kumbhakar, S. C.: 1987, The specification of technical and allocative inefficiency in stochastic production and profit frontiers, *Journal of Econometrics* **34**(3), 335 – 348.

  URL: *http://www.sciencedirect.com/science/article/pii/0304407687900169*

Kumbhakar, S. C.: 1990, Production frontiers, panel data, and time-varying technical inefficiency, *Journal of Econometrics* **46**(1 - 2), 201–211.

  URL: *http://www.sciencedirect.com/science/article/pii/030440769090055X*

Kumbhakar, S. C.: 1997, Modeling allocative inefficiency in a translog cost function and cost share equations: an exact relationship, *Journal of Econometrics* **76**(1), 351–356.

Kumbhakar, S. C. and Tsionas, E. G.: 2005, Measuring technical and allocative inefficiency in the translog cost system: a Bayesian approach, *Journal of Econometrics* **126**(2), 355–384. <ce:title>Current developments in productivity and efficiency measurement</ce:title>.

  URL: *http://www.sciencedirect.com/science/article/pii/S0304407604001162*

Kumbhakar, S. and Lovell, C.: 2000, *Stochastic Frontier Analysis*, 1st edn, Cambridge University Press, United Kingdom.

Kurkalova, L. A. and Carriquiry, A.: 2003, Input- and output-oriented technical efficiency of ukrainian collective farms, 1989-1992: Bayesian analysis of a stochastic production frontier model, *Journal of Productivity Analysis* **20**(2), 191–211.

Kuroda, Y. and Yotopoulos, P. A.: 1978, A microeconomic analysis of production behavior of the farm household in japan - a profit function approach, *The Economic Review (Japan)* **29**(2), 116–129.

Land, K. C., Lovell, C. A. K. and Thore, S.: 1993, Chance-constrained data envelopment analysis, *Managerial and Decision Economics* **14**(6), 541–554.

Lau, L. J., Lin, W.-L. and Yotopoulos, P. A.: 1978, The linear logarithmic expenditure system: An application to consumption-leisure choice, *Econometrica* **46**(4), 843–868.
URL: *http://www.jstor.org/stable/1909753*

Lau, L. J. and Yotopoulos, P. A.: 1971, A test for relative efficiency and application to indian agriculture, *The American Economic Review* **61**(1), 94–109.
URL: *http://www.jstor.org/stable/1910544*

Lee, Y. and Schmidt, P.: 1993, *The Measurement of Productive Efficiency : Techniques and Application*, Oxford University Press, UK, chapter 8, pp. 237 – 255.

Leibenstein, H.: 1966, Allocative efficiency vs. "x-efficiency", *The American Economic Review* **56**(3), 392–415.

Leibenstein, H.: 1973, Competition and x-efficiency: Reply, *Journal of Political Economy* **81**(3), pp.765–777.

Leibenstein, H.: 1978, *General X-efficiency theory and economic development*, Oxford University Press.

Leibowitz, A. and Friedman, B.: 1979, Family bequests and the derived demand for health inputs, *Economic inquiry* **17**(3), 419.

Lesaffre, E. and Lawson, A. B.: 2012, *Bayesian biostatistics*, John Wiley & Sons.

Lindley, D. V. and Smith, A. F. M.: 1972, Bayes estimates for the linear model, *Journal of the Royal Statistical Society. Series B (Methodological)* **34**(1), pp.1–41.
URL: *http://www.jstor.org/stable/2985048*

Lipton, M.: 1968, The theory of optimising peasant, *Journal of Development Studies* **4**(3), 26.

Liu, C., Liu, J. and Rubin, D.: 1993, A control variable for assessment the convergence of the gibbs sampler, *Proceedings of the Statistical Computing Section of the American Statistical Association*, pp. 74–78.

Lofgren, H. and Robinson, S.: 1999, Nonseparable farm household decisions in a computable general equilibrium model, *American Journal of Agricultural Economics* **81**(3), 663–670.

Lopez, R. E.: 1982, Applications of duality theory to agriculture, *Western Journal of Agricultural Economics* **7**(02), 353–366.
URL: *http://purl.umn.edu/32280*

Lopez, R. E.: 1984, Estimating labor supply and production decisions of self-employed farm producers, *European Economic Review* **24**(1), 61 – 82.
URL: *http://www.sciencedirect.com/science/article/pii/0014292184900138*

Lopez, R. E.: 1986, *Structural Models of the Farm Household That Allow for Interdependent Utility and Profit Maximization Decisions*, The John Hopkins University Press, chapter 11, pp. 306 – 325.

Loureiro, M. L.: 2009, Farmers' health and agricultural productivity, *Agricultural Economics* **40**(4), 381–388.

Lovell, C.: 1993, *Production Frontiers and Productive Efficiency. In The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press, chapter 1, pp. 3–67.

Maceachern, S. N. and Berliner, L. M.: 1994, Subsampling the gibbs sampler, *The American Statistician* **48**(3), 188–190.

Malaney, P., Spielman, A. and Sachs, J.: 2004, The malaria gap, *The American Journal of Tropical Medicine and Hygiene* **71**(2 suppl), 141–146.

Malaria Atlas Project: 2011, Posterior distribution of candidate map at different raster value.
URL: *www.map.ox.ac.uk*

Mattila-Wiro, P.: 1999, *Economic theories of the household: a critical review*, number 159, United Nations University, World Institute for Development Economics Research.

McCarthy, D., Wolf, H. and Wu, Y.: 2000, The growth costs of malaria, National Bureau of Economic Research.

URL: *http://www.nber.org/papers/w7541*

McCulloch, R. and Rossi, P. E.: 1994, An exact likelihood analysis of the multinomial probit model, *Journal of Econometrics* **64**(1-2), 207–240.

McElroy, M. B. and Horney, M. J.: 1981, Nash-bargained household decisions: Toward a generalization of the theory of demand, *International Economic Review* **22**(2), pp.333–349.

URL: *http://www.jstor.org/stable/2526280*

Meeusen, W. and van den Broeck, J.: 1977, Efficiency estimation from cobb-douglas production functions with composed error, *International Economic Review* **18**(2), 435–444.

URL: *http://www.jstor.org/stable/2525757*

Mekonnen, A., Brännlund, R., Kriström, B., Nilsagård, H. et al.: 1999, Rural household biomass fuel production and consumption in ethiopia: a case study., *Special issue: the expanding horizons of forest economics. Selected papers from the Forest Resource Economics meeting, September 8-12, 1997, Umeå, Sweden.*, Vol. 5, pp. 69–97.

Melfi, C. A.: 1984, Estimation and decomposition of productive efficiency in a panel data model: an application to electric utilities.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E.: 1953, Equation of state calculations by fast computing machines, *The Journal of Chemical Physics* **21**, 1087.

Michelena, A.: 1965, *Culture Change, Applied Anthropology and Economic Development*, Vol. 67, University of California Press.

Micklewright, J.: 1989, Towards a household model of {UK} domestic energy demand, *Energy Policy* **17**(3), 264 – 276.

URL: *http://www.sciencedirect.com/science/article/pii/0301421589900517*

Mills, A.: 1992, The economic evaluation of malaria control technologies: The case of nepal, *Soc.Sci.Med.* **34**(9), 965 – 972.

Mincer, J. et al.: 1963, Market prices, opportunity costs, and income effects, *Measurement in economics* pp. 67–82.

Mochebelele, M. T. and Winter-Nelson, A.: 2000, Migrant labor and farm technical efficiency in lesotho, *World Development* **28**(1), 143–153.
URL: *http://www.sciencedirect.com/science/article/pii/S0305750X99001163*

Mood, A., Graybill, F. and Boes, D.: 1974, *Introduction to the Theory of Statistics*, third edn, McGraw-Hill, New York.

Mrdonn: 2012, African geography.
URL: *http://africa.mrdonn.org/geography.html*

Murillo-Zamorano, L. R.: 2004, Economic efficiency and frontier techniques, *Journal of Economic Surveys* **18**(1), 33–77.

Myers, J.: 1988, The value of ideal contigency markets in agriculture, *American Journal of Agricultural Economics* **70**(2), 255 – 267.

Mykland, P., Tierney, L. and Yu, B.: 1995, Regeneration in markov chain samplers, *Journal of the American Statistical Association* **90**(429), 233–241.

NBS Nigeria: 2011, General household panel survey (post planting 2010) basic information document, *Technical report*, National Bureau of Statiatics.

Newbery, D. M. and Stiglitz, J. E.: 1981, The theory of stabilization: A study in the economics of risk.

Newey, W. K. and West, K. D.: 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**(3), 703–708.
URL: *http://www.jstor.org/stable/1913610*

Nigeria Ministry of Health: 2008, A road map for malaria control in nigeria: Strategic plan 2009 - 2013, *Technical report*, National Malaria Control Programme, Abuja, Nigeria.

Nummelin, E.: 2004, *General Irreducible Markov Chains and Non-Negative Operators*, Cambridge Tracts in Mathematics, Cambridge University Press.

Nur, D., Mengersen, K. and Wolff, R.: 2005, Phase randomization: A convergence diagnostic test for mcmc, *Australian & New Zealand Journal of Statistics* **47**(3), 309–323.

Nur, E.: 1993, The impact of malaria on labour use and efficiency in the sudan, *Social Science and Medicine* **37**(9), 1115–1119.

Nylander, J. A., Wilgenbusch, J. C., Warren, D. L. and Swofford, D. L.: 2008, Awty (are we there yet?): a system for graphical exploration of mcmc convergence in bayesian phylogenetics, *Bioinformatics* **24**(4), 581–583.

O'Donnell, C.: 2012, Econometric estimation of distance functions and associated measures of productivity and efficiency change, *Journal of Productivity Analysis* pp. 1–14.

O'Donnell, C. J. and Griffiths, W. E.: 2006, Estimating state-contingent production frontiers, *American Journal of Agricultural Economics* **88**(1), 249–266.

O'Donnell, J. and Coelli, T.: 2005, A Bayesian approach to imposing curvature on distance functions, *Journal of Econometrics* **126**, 493–523.

Ogundele, O. and Okoruwa, V.: 2006, Technical efficiency differentials in rice production technologies in nigeria, *Research Paper 154*, AERC, Nairobi.
URL: *http://www.aercafrica.org/documents/rp154.pdf*

Olarinde, L., Ajao, A. and Okunola, S.: 2008, Determinants of technical efficiency in bee-keeping farms in oyo state, nigeria: A stochastic production frontier approach, *Research Journal of Agriculture and Biological Sciences* **4**(1), 65–69.

Olesen, O. B. and Petersen, N. C.: 1995, Chance constrained efficiency evaluation, *Management Science* **41**(3), 442–457.

Olson, J. A., Schmidt, P. and Waldman, D. M.: 1980, A Monte Carlo study of estimators of stochastic frontier production functions, *Journal of Econometrics* **13**(1), 67–82.
URL: *http://www.sciencedirect.com/science/article/pii/0304407680900433*

Omamo, S. W.: 1998, Farm to market transaction costs and specialisation in small scale agriculture: Explorations with a non-separable household model, *The Journal of Development Studies* **35**(2), 152–163.

Onwujekwe, O., Hanson, K. and Fox-Rushby, J.: 2004, Inequalities in purchase of mosquito nets and willingness to pay for insecticide-treated nets in nigeria: Challenges for malaria control interventions, *Malaria Journal* **3**(1), 6.

Onwujekwe, O., Hanson, K. and Fox-Rushby, J.: 2005, Do divergences between stated and actual willingness to pay signify the existence of bias in contingent valuation surveys?, *Soc Sci Med* **60**, 525–536.

Owen, J.: 1964, *The Supply of labour and the Demand for Recreation*, PhD thesis, Columbia University.

Patil, A. P., Gething, P. W., Piel, F. B. and Hay, S. I.: 2011, Bayesian geostatistics in health cartography: the perspective of malaria, *Trends in Parasitology* **27**(6), 246–253.

Paul, R., MacEachern, S. N. and Berliner, L. M.: 2012, Assessing convergence and mixing of mcmc implementations via stratification, *Journal of Computational and Graphical Statistics* **21**(3), 693–712.

Philippe, A. and Robert, C. P.: 2001, Riemann sums for mcmc estimation and convergence monitoring, *Statistics and Computing* **11**(2), 103–115.

Pitt, M. M. and Lee, L.-F.: 1981, The measurement and sources of technical inefficiency in the indonesian weaving industry, *Journal of Development Economics* **9**(1), 43 – 64.
  URL: *http://www.sciencedirect.com/science/article/pii/0304387881900043*

Pitt, M. M. and Rosenzweig, M. R.: 1985, Health and nutrient consumption across and within farm households, *The Review of Economics and Statistics* **67**(2), 212–223.
  URL: *http://www.jstor.org/stable/1924720*

Polanyi, K.: 1944, *The Great Transformation: The Political and Economic Origins of Our Time*, 1st edn, Beacon in Boston. Originally published: New York: Farrar & Rinehart, 1944 and reprinted in 1957 by Beacon in Boston.

Politis, D. N. and Romano, J. P.: 1994, Large sample confidence regions based on sub-samples under minimal assumptions, *The Annals of Statistics* **22**(4), 2031–2050.

Politis, D., Romano, J. P. and Wolf, M.: 1997, Subsampling for heteroskedastic time series, *Journal of Econometrics* **81**(2), 281 – 317.

Pollak, R. A. and Wachter, M. L.: 1975, The relevance of the household production function and its implications for the allocation of time, *Journal of Political Economy* **83**(2), pp. 255–278.
  URL: *http://www.jstor.org/stable/1830922*

Pollak, R. A. and Wachter, M. L.: 1977, Reply: "pollak and wachter on the household production approach", *Journal of Political Economy* **85**(5), 1083–1086.
  URL: *http://www.jstor.org/stable/1830348*

Polson, N. G.: 1993, *Convergence of Markov chain Monte Carlo algorithms*, University of Chicago, Graduate School of Business, Department of Economics.

Propp, J. G. and Wilson, D. B.: 1996, Exact sampling with coupled markov chains and applications to statistical mechanics, *Random Structures & Algorithms* **9**(1-2), 223–252.

Raftery, A. E., Lewis, S. et al.: 1992, How many iterations in the gibbs sampler, *Bayesian statistics* **4**(2), 763–773.

Reardon, T.: 1997, Using evidence of household income diversification to inform study of the rural nonfarm labor market in africa, *World Development* **25**(5), 735 – 747.
  URL: *http://www.sciencedirect.com/science/article/pii/S0305750X96001374*

Rees, A.: 1979, *The Economics of Work and Pay*, 2 edn, Harper and Row. Harper International Edition.

Reinhard, S., Lovell, C. A. K. and Thijssen, G.: 1999, Econometric estimation of technical and environmental efficiency: An application to dutch dairy farms, *American Journal of Agricultural Economics* **81**(1), 44–60.
  URL: *http://www.jstor.org/stable/1244449*

Reinhard, S., Lovell, C. K. and Thijssen, G. J.: 2000, Environmental efficiency with multiple environmentally detrimental variables; estimated with sfa and dea, *European Journal of Operational Research* **121**(2), 287–303.

Renkow, M.: 1990, Household inventories and marketed surplus in semisubsistence agriculture, *American Journal of Agricultural Economics* **72**(3), 664–675.
URL: *http://www.jstor.org/stable/1243036*

Richmond, J.: 1974, Estimating the efficiency of production, *International Economic Review* **15**(2), 515–521.
URL: *http://www.jstor.org/stable/2525875*

Rios, A. R. and Shively, G.: 2005, Farm size and non-parametric efficiency measurements for coffee farms in vietnam, Univ. of Minnesota(umn).

Ripley, B.: 1987, *Stochastic Simulation*, John Wiley & Sons Ltd.

Ritter, C. and Simar, L.: 1997, Pitfalls of normal-gamma stochastic frontier models, *Journal of Productivity Analysis* **8**(2), 167–182.

Ritter, C. and Tanner, M. A.: 1992, Facilitating the gibbs sampler: The gibbs stopper and the griddy-gibbs sampler, *Journal of the American Statistical Association* **87**(419), 861–868.

Robert, C. and Casella, G.: 2004, *Monte Carlo Statistical Methods*, 2 edn, Springer-Verlag New York.

Robert, C. P.: 1995, Convergence control methods for markov chain monte carlo algorithms, *Statistical Science* **10**(3), 231–253.

Roberts, G.: 1992, Convergence diagnostics of the gibbs sampler, *Bayesian statistics* **4**, 777–784.

Roberts, G.: 1996, Methods for estimating l2 convergence of markov chain-monte carlo techniques, *Bayesian Analysis in Statistics and Econometrics: Essays in Honor of Arnold Zellner* **309**, 373.

Rodriguez-Alvarez, A. and Knox Lovell, C.: 2004, Excess capacity and expense preference behaviour in national health systems: An application to the spanish public hospitals, *Health Economics* **13**(2), 157–169.

Roll, S. M. Y.: 1985, Solving for "x": Theory and measurement of allocative and x-efficiency at the plant level, *Journal of Behavioral Economics* **14**.

Rosenthal, J. S.: 1993, Rates of convergence for data augmentation on finite sample spaces, *The Annals of Applied Probability* **3**(3), 819–839.
  URL: *http://www.jstor.org/stable/2959692*

Rosenthal, J. S.: 1995a, Convergence rates for markov chains, *Siam Review* **37**(3), 387–405.

Rosenthal, J. S.: 1995b, Minorization conditions and convergence rates for markov chain monte carlo, *Journal of the American Statistical Association* **90**(430), 558–566.
  URL: *http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1995.10476548*

Rosenzweig, M. R. and Schultz, T. P.: 1983, Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight, *Journal of Political Economy* **91**(5), 723–746.
  URL: *http://www.jstor.org/stable/1837367*

Ruben, R. and van Ruijven, A.: 2001, Technical coefficients for bio-economic farm household models: A meta-modelling approach with applications for southern mali, *Ecological Economics* **36**(3), 427 – 441.
  URL: *http://www.sciencedirect.com/science/article/pii/S0921800900002408*

Russell, S.: 2004, The economic burden of illness for households in developing countries: A review of studies focusing on malaria, tuberculosis, and human immunodeficiency virus/acquired immunodeficiency syndrome, *The American Journal of Tropical Medicine and Hygiene* **71**(2 suppl), 147–155.

Sachs, J. and Malaney, P.: 2002, The economic and social burden of malaria, *Nature*

415(6872), 680 – 685.

URL: *http://dx.doi.org/10.1038/415680a*

Saha, A.: 1994, A two-season agricultural household model of output and price uncertainty, *Journal of Development Economics* **45**(2), 245–269.

Saha, A. and Stroud, J.: 1994, A household model of on-farm storage under price risk, *American Journal of Agricultural Economics* **76**(3), 522–534.

Samba, E.: 2001, The malaria burden and africa, *Am J Trop Med Hyg* **64**(1), ii–iii.

URL: *http://www.ajtmh.org*

Sauer, J. and Mendoza-Esclante, A.: 2007, Poor but allocatively efficient - evidence from the eastern amazon, *Agricultural Economics* **37**(1), 113–126.

Schmidt, P.: 1985, Frontier production functions, *Econometric Reviews* **4**(2), 289–328.

Schmidt, P. and Knox Lovell, C.: 1979, Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers, *Journal of econometrics* **9**(3), 343–366.

Schmidt, P. and Lovell, C.: 1980, Estimating stochastic production and cost frontiers when technical and allocative inefficiency are correlated, *Journal of Econometrics* **13**(1), 83–100.

Schmidt, P. and Sickles, R. C.: 1984, Production frontiers and panel data, *Journal of Business & Economic Statistics* **2**(4), 367–374.

URL: *http://www.jstor.org/stable/1391278*

Schruben, L., Singh, H. and Tierney, L.: 1983, Optimal tests for initialization bias in simulation output, *Operations Research* **31**(6), 1167–1178.

URL: *http://www.jstor.org/stable/170843*

Schruben, L. W.: 1982, Detecting initialization bias in simulation output, *Operations Research* **30**(3), 569–590.

Schultz, T.: 1964, *Transforming Traditional Agriculture.*, New Haven: Yale University Press.

Schwartzman, D.: 1973, Competition and efficiency: Comment, *Journal of Political Economy* **81**(3), 756–764.

   URL: *http://www.jstor.org/stable/1831039*

Seitz, W. D.: 1970, The measurement of efficiency relative to a frontier production function, *American Journal of Agricultural Economics* **52**(4), pp. 505–511.

   URL: *http://www.jstor.org/stable/1237257*

Seyoum, E. T., Battese, G. E. and Fleming, E. M.: 1998, Technical efficiency and productivity of maize producers in eastern ethiopia: a study of farmers within and outside the sasakawa-global 2000 project, *Agricultural Economics* **19**(3), 341–348.

   URL: *http://www.sciencedirect.com/science/article/pii/S0169515098000371*

Seyoum Taffesse, A., Dorosh, P. and Asrat, S.: 2011, Crop production in ethiopia: regional patterns and trends.

Shah, M., Farman, A. and Humayun, K.: 1994, technical efficiency of major crops in the north-west frontier province of pakistan, *Sarhad Journal of Agriculture* **X**(6).

Shapiro, K. H.: 1983, Efficiency differentials in peasant agriculture and their implications for development policies, *Journal of Development Studies* **19**(2), 179.

Shapiro, K. H. and Mãller, J.: 1977, Sources of technical efficiency: The roles of modernization and information, *Economic Development and Cultural Change* **25**(2), 293–310.

   URL: *http://www.jstor.org/stable/1152862*

Sickles, R. C., Good, D. and Johnson, R. L.: 1986, Allocative distortions and the regulatory transition of the us airline industry, *Journal of Econometrics* **33**(1), 143–163.

Silverman, B. W.: 1986, *Density estimation for statistics and data analysis*, Vol. 26, CRC press.

Simar, L.: 1992, Estimating efficiencies from frontier models with panel data: A comparison of parametric, non-parametric and semi-parametric methods with bootstrapping, *Journal of Productivity Analysis* **3**(1-2), 171–203.

Simtowe, F. and Zeller, M.: 2006, The impact of access to credit on the adoption of hybrid maize in malawi: An empirical test of an agricultural household model under credit market failure.

Singh, I., Squire, L. and Strauss, J.: 1986, *Agricultural Household Models - Extension, Applications and Policy*, For the World Bank: The John Hopkins University Press, Baltimore and London.

Skoufias, E.: 1994, Using shadow wages to estimate labor supply of agricultural households, *American Journal of Agricultural Economics* **76**(2), 215.

Smith, A. F. M. and Gelfand, A. E.: 1992, Bayesian statistics without tears: A sampling-resampling perspective, *The American Statistician* **46**(2), 84–88.

Sofola, Y.: 2008, The burden of malaria on nigeria, *Punch Newspaper* .

Stevenson, R.: 1980, Likelihood functions for generalized stochastic frontier estimation, *Journal of Econometrics* **13**, 57–66.

Stigler, G. J.: 1976, Association the xistence of x-efficiency, *The American Economic Review* **66**(1), 213–216.
  URL: *http://www.jstor.org/stable/1804963* .

Strauss, J.: 1984, Marketed surpluses of agricultural households in sierra leone, *American Journal of Agricultural Economics* **66**(3), 321–331.
  URL: *http://www.jstor.org/stable/1240799*

Strauss, J.: 1986, Does better nutrition raise farm productivity?, *Journal of Political Economy* **94**(2), 297–320.
  URL: *http://www.jstor.org/stable/1837406*

Swamy, P. A. V. B.: 1970, Efficient inference in a random coefficient regression model, *Econometrica* **38**(2), 311–323.
  URL: *http://www.jstor.org/stable/1913012*

Tadesse, B. and Krishnamoorthy, S.: 1997, Technical efficiency in paddy farms of tamil nadu: An analysis based on farm size and ecological zone, *Agricultural Economics*

16(3), 185–192.

URL: *http://www.sciencedirect.com/science/article/pii/S0169515097000042*

Tanner, M. A. and Wong, W. H.: 1987, The calculation of posterior distributions by data augmentation, *Journal of the American Statistical Association* **82**(398), 528–540.

URL: *http://www.jstor.org/stable/2289457*

Tanzania Ministry of Health and Social Welfare: 2008, Medium term malaria strategic plan 2008 - 2013, *Technical report*, Tanzania National Malaria Control Programme.

URL: *natnets.org/attachments/article/65/MTMSP202008-2013.pdf*

Tanzania National Bureau of Statistics: 2009, Basic information document - national panel survey (nps 2008-2009), *Technical report*, Tanzania National Bureau of Statistics.

Taylor, J. E. and Adelman, I.: 2003, Agricultural household models: Genesis, evolution, and extensions, *Review of Economics of the Household* **1**(1-2), 33–58.

Taylor, T. G., Drummond, H. E. and Gomes, A. T.: 1986, Agricultural credit programs and production efficiency: An analysis of traditional farming in southeastern minas gerais, brazil, *American Journal of Agricultural Economics* **68**(1), 110–119.

URL: *http://www.jstor.org/stable/1241655*

Taylor, T. G. and Scott Shonkwiler, J.: 1986, Alternative stochastic specifications of the frontier production function in the analysis of agricultural credit programs and technical efficiency, *Journal of Development Economics* **21**(1), 149–160.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. and Farmer, J. D.: 1992, Testing for nonlinearity in time series: the method of surrogate data, *Physica D: Nonlinear Phenomena* **58**(1), 77 – 94.

Thijssen, G.: 1988, Estimating a labour supply function of farm households, *European Review of Agricultural Economics* **15**(1), 67–78.

URL: *http://erae.oxfordjournals.org/content/15/1/67.abstract*

Thorner, D., Kerblay, B. and Smith, R.: 1966, *A.V.Chayanov on the Theory of Peasant Economy*, 1st edn, Richard D. Irwin Inc, Homewood, Illinois.

Timmer, C. P.: 1971, Using a probabilistic frontier production function to measure technical efficiency, *Journal of Political Economy* **79**(4), 776–794.
URL: *http://www.jstor.org/stable/1830828*

Tomberlin, D. and Holloway, G.: 2008, Bayesian hierarchical estimationof technical efficiency in a fishery, *Applied Economics Letters* **17**(2), 201–204. doi:.

Tsionas, E. G.: 2002, Stochastic frontier models with random coefficients, *Journal of Applied Econometrics* **17**(2), 127 – 147.

Udry, C.: 1996, Gender, agricultural production, and the theory of the household, *Journal of Political Economy* **104**(5), 1010–1046.
URL: *http://www.jstor.org/stable/2138950*

Uguru, N. P., Onwujekwe, O., Uzochukwu, B., Igiliegbe, G. and Eze, S.: 2009, Inequities in incidence, morbidity and expenditures on prevention and treatment of malaria in southeast nigeria, *BMC international health and human rights* **9**(1).

Ulimwengu, J.: 2009, Farmers health and agricultural productivity in rural ethiopia, *African Journal of Agricultural and Resource Economics* **3**(2), 83–100.

Ulimwengu, J. M.: 2008, Farmers health and agricultural efficiency in rural ethiopia: A stochastic production frontier approach.

UNEP: 2007, Africa economic outlook 2.
URL: *http://www.eoearth.org/article*

University of North Carolina at Chapel Hill and Bauer, P.: 1985, *An Analysis of Multiproduct Technology and Efficiency Using the Joint Cost Function and Panel Data: An Application to US Airline Industry*, University Microfilms.

van den Broeck, J., Koop, G., Osiewalski, J. and Steel, M. F. J.: 1994, Stochastic frontier models: A bayesian perspective, *Journal of Econometrics* **61**(2), 273–303.
URL: *http://www.sciencedirect.com/science/article/pii/0304407694900876*

Van Der Vaart, A. W. and Wellner, J. A.: 1996, *Weak Convergence*, Springer.

VanDerwerken, D. N.: 2015, *Monitoring and Improving Markov Chain Monte Carlo Convergence by Partitioning*, PhD thesis, Duke University, Department of Statistical Science.

Varian, H. R.: 1985, Non-parametric analysis of optimizing behavior with measurement error, *Journal of Econometrics* **30**(1-2), 445–458.

Wadud, A. and White, B.: 2000, Farm household efficiency in bangladesh: a comparison of stochastic frontier and dea methods, *Applied Economics* **32**(13), 1665–1673. doi:.

Waldman, D. M.: 1984, Properties of technical efficiency estimators in the stochastic frontier model, *Journal of Econometrics* **25**(3), 353–364.
URL: *http://www.sciencedirect.com/science/article/pii/030440768490006X*

Wales, T. J. and Woodland, A. D.: 1976, Estimation of household utility functions and labor supply response, *International Economic Review* **17**(2), 397–410.

Wang, H. and Schmidt, P.: 2002, One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels, *Journal of Productivity Analysis* **18**(2), 129–144.

Webb, J. L. A.: 2011, Malaria in africa, *History Compass* **9**(3), 162–170.

Wei, G. C. G. and Tanner, M. A.: 1990, A monte carlo implementation of the em algorithm and the poor man's data augmentation algorithms, *Journal of the American Statistical Association* **85**(411), 699–704.

Welsch, D. E.: 1965, Response to economic incentive by abakaliki rice farmers in eastern nigeria, *Journal of Farm Economics* **47**(4), 900–914.
URL: *http://www.jstor.org/stable/1236333*

WHO: 2012, Progress and impact series: Focus on nigeria, *Technical Report 4*, World Health Organisation.

Winsten, C.: 1957, Discussion on mr. farrell's paper, *Journal of the Royal Statistical Society* **120**, 282–284.

Wiser, M.: 2010, *Protozoa and Human Disease*, Taylor and Francis.

   URL: *http://www.tulane.edu/wiser/malaria*

Wouterse, F.: 2010, Migration and technical efficiency in cereal production: evidence from burkina faso, *Agricultural Economics* **41**(5), 385–395.

www.mapsofworld.com: 2013, Africa.

   URL: *http://www.mapsofworld.com/africa/maps/africa-map.gif*

www.vtaide.com: 2013, The life cycle of malaria.

   URL: *http://www.vtaide.com/png/images/mosquito-cycle.jpg*

Yotopoulos, P.A, L. L. J. and Lin, W.-L.: 1976, Microeconomic output supply and factor demand functions in the agriculture of the province of taiwan, *American Journal of Agricultural Economics* **58**(2), 333–340.

Yu, B.: 1995, [bayesian computation and stochastic systems]: Comment: Extracting more diagnostic information from a single run using cusum path plot, *Statist. Sci.* **10**(1), 54–58.

Yu, B. and Mykland, P.: 1998, Looking at markov samplers through cusum path plots: A simple diagnostic idea, *Statistics and Computing* **8**(3), 275–286.

Yu, B. et al.: 1995, Estimating l1 error of kernel estimator: monitoring convergence of markov samplers, *Technical report*, Technical report, Dept. of Statistics, University of California, Berkeley.

Zellner, A.: 1971, *An Introduction to Bayesian Inference in Econometrics*, wiley classics library edn, John Wiley & Sons, INC., New York.

Zellner, A. and Min, C.-K.: 1995, Gibbs sampler convergence criteria, *Journal of the American Statistical Association* **90**(431), 921–927.

   URL: *http://www.jstor.org/stable/2291326*

A. Malaria Coefficient Pdf For Each Country

Figure A.1.: Malaria Coefficient pdf for Nigeria
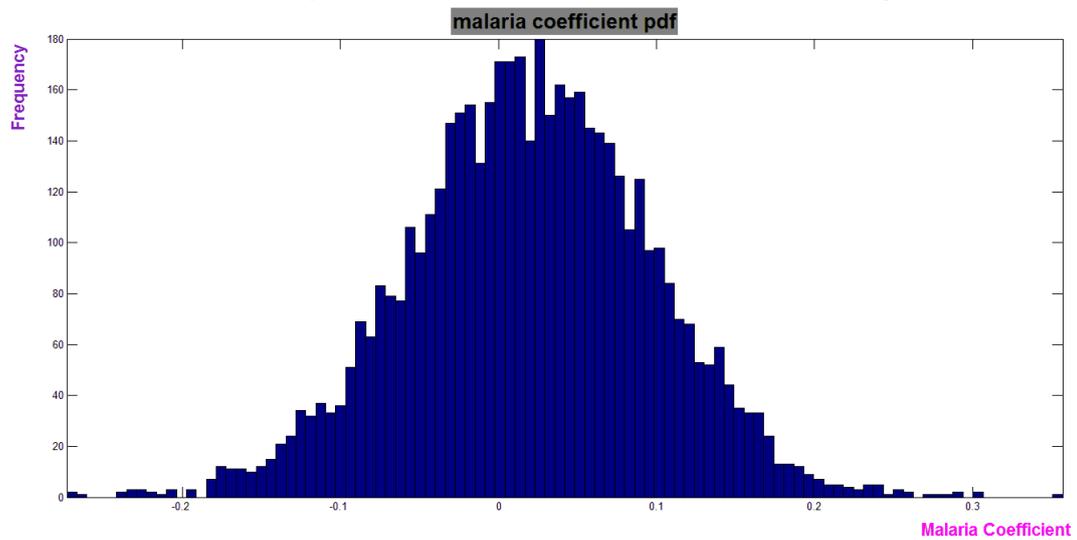
Figure A.2.: Malaria Coefficient pdf for Ethiopia
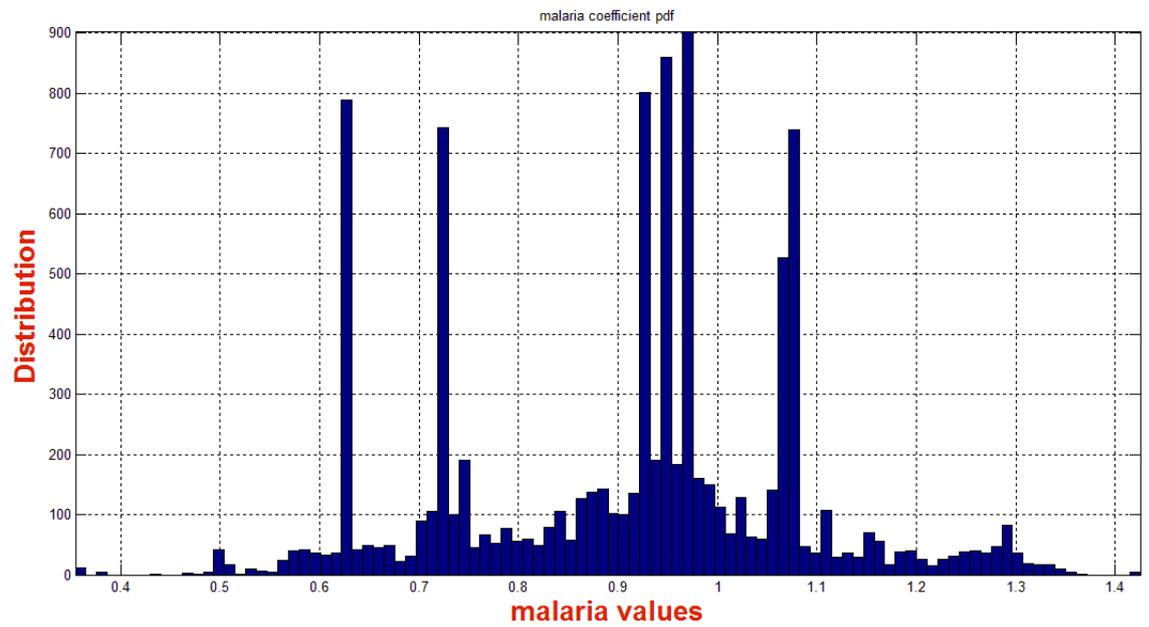
Figure A.3.: Malaria Coefficient pdf for Tanzania
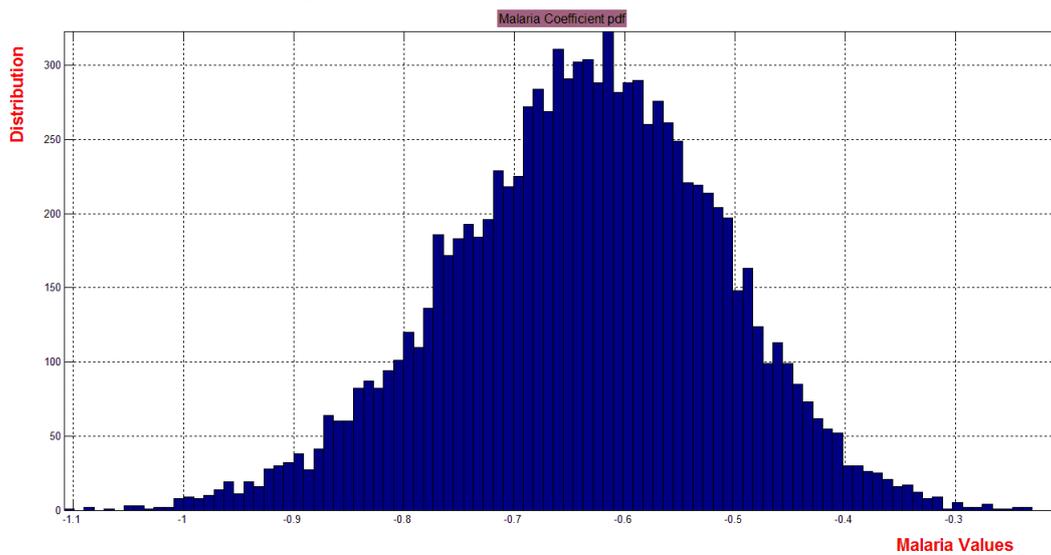
B. Posterior Prediction For Each Country

Figure B.1.: Nigeria Posterior Prediction
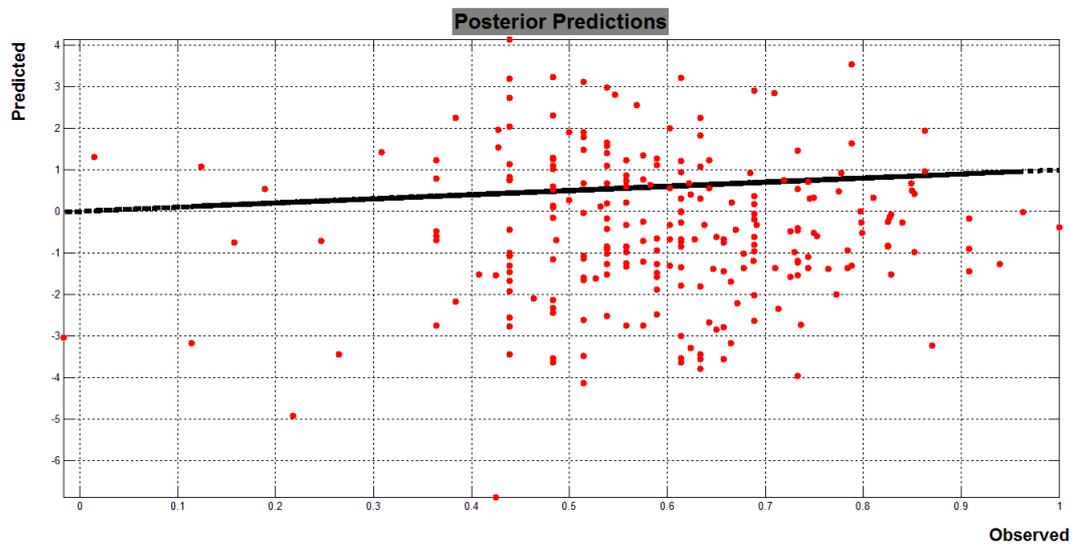
Figure B.2.: Ethiopia Posterior Prediction
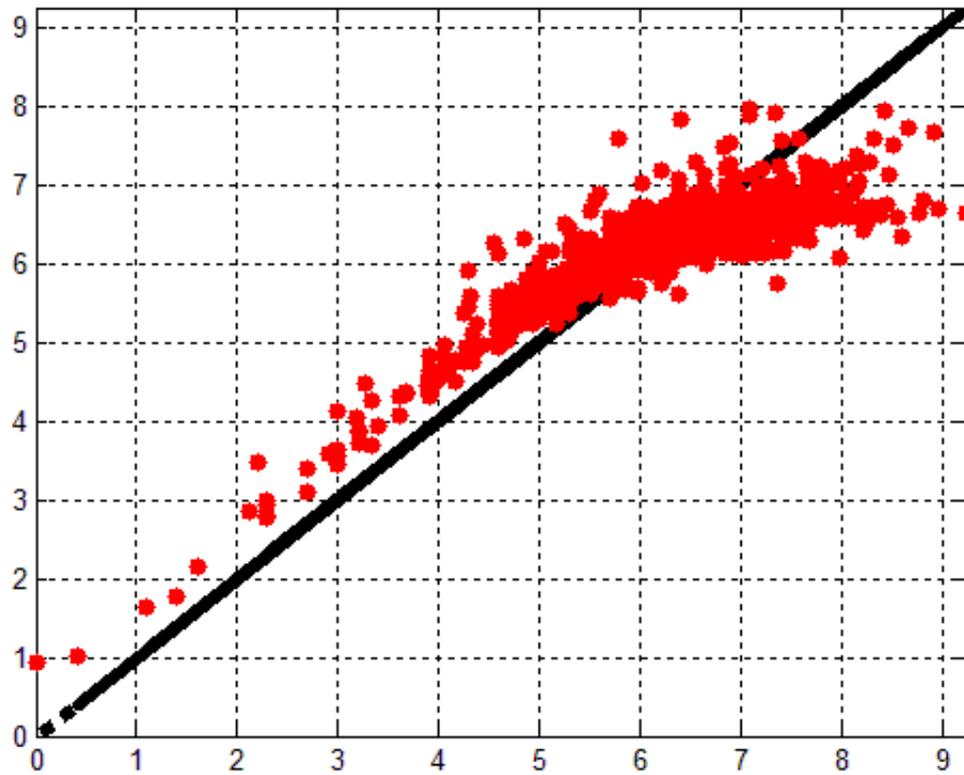
Figure B.3.: Tanzania Posterior Prediction
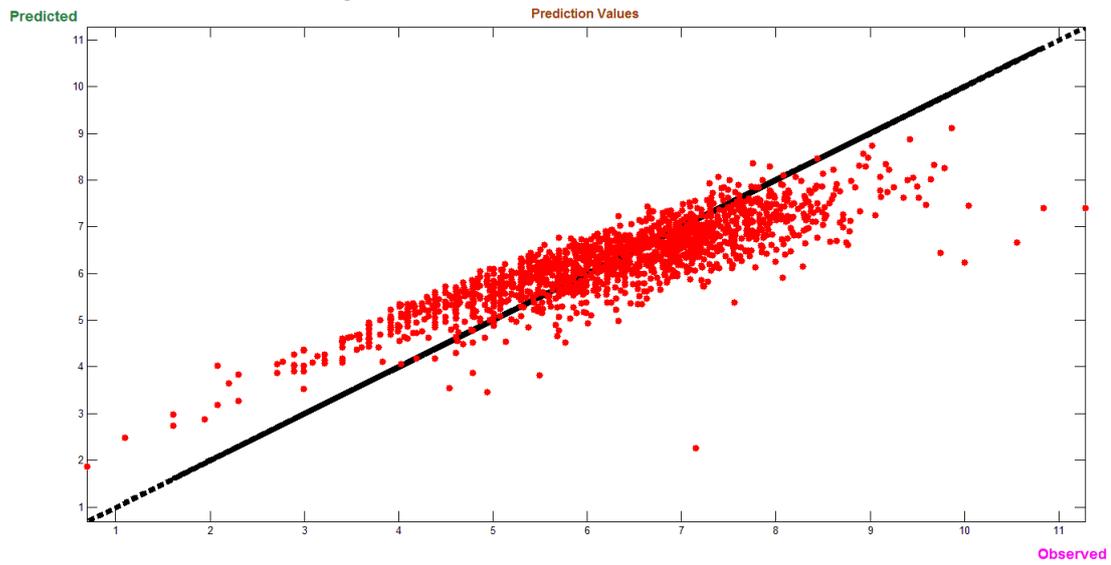
C. Markov chain Monte Carlo Convergence Diagnostics

C.1. The Graphical Approach

C.1.1. The CuSum Plot Method

In using the CuSum plot, the researcher needs to understand the difference between the smoothness and "hairiness" of a plot. A plot is smooth when it is formed from line segments with the same or similar slopes while a "hairy" plot is formed from line segments which alternates between positive and negative slope such that each point corresponds to a local optimum (minimum or maximum) point. In other words, you get the "hairiness" of a plot by counting the number of optimum points present in the plot.

If we define

$$
d_t = \begin{cases} 1 \ \ if \ \ S_{T-1} > S_T \ andS_T < S_{T+1} \\ \quad or \ \ S_{T-1} < S_T \ and \ S_T > S_{T+1} \\ 0 \qquad\qquad else \end{cases}
\tag{C.1}
$$

for all $t = n_0 + 1, ..., n - 1$. Then,

$$
D_{n_0,n} = \frac{1}{n - n_0} \sum_{t=n_0+1}^{n-1} d_t
\tag{C.2}
$$

takes values 0 and 1, where a value of 0 indicates a completely smooth plot and a value of 1, means maximum 'hairiness'.

Brooks and Roberts (1998) suggests a boundary in which convergence is expected to occur and should lie within the bounds:

$$
\frac{1}{2} \pm Z_{\alpha/2} \sqrt{\frac{1}{4(n - n_0)}}
\tag{C.3}
$$

$100(1-\alpha/2)\%$ of the time.

C.2. The Formal Diagnostic Approach

C.2.1. Variance Ratio Methods

The scale parameter say $\boldsymbol{\theta}(\boldsymbol{x})$ involves the between-chain variance B/n also denoted by $\bar{\theta}_i$, and the within-chain variance, s_i^2 with mean W. Brooks and Roberts (1998) write this as:

$$\frac{B}{n} = \frac{1}{m-1} \sum_{i=1}^{m} (\bar{\theta}_i - \bar{\theta})^2 \tag{C.4}$$

where $\bar{\theta}_i = \frac{1}{n} \sum_{t=n+1}^{2n} \theta_i^t; \qquad \bar{\theta} = \frac{1}{m} \sum_{i=1}^{m} \bar{\theta}_i$

The quantity $\theta_i^t = \theta(\boldsymbol{x}_i^t)$ is the tth observation of θ from chain i. We calculate W, the mean of the m within-sequence variances, s_i^2 with $n-1$ degrees of freedom as:

$$W = \frac{1}{m} \sum_{i=1}^{m} s_i^2 \tag{C.5}$$

where $s_i^2 = \frac{1}{n-1} \sum_{t=n+1}^{2n} (\theta_i^t - \bar{\theta}_i)^2$

The variance, σ^2, of the stationary (target) distribution is estimated by:

$$\hat{\sigma}^2 = \frac{n-1}{n} W + \frac{1}{n} B \tag{C.6}$$

Equation (D.6) overestimates σ^2 if the starting distribution is overdispersed but it becomes unbiased when stationarity is attained in the distribution, which is when $n \to \infty$. Gelman and Rubin (1992) propose a factor $\hat{R}_c$ by which convergence can be monitored. This is defined as:

$$\hat{R}_c = \frac{d+3}{d+1} \frac{\hat{\sigma}^2}{W} \tag{C.7}$$

Gelman and Rubin (1992) refer to this factor as the *Potential Scale Reduction Factor* (PSRF). The factor reduces to 1 as $n \to \infty$, in other words, the larger the simulation the more likely it reaches the stationary distribution.

Brooks and Gelman (1998) generalise the Gelman and Rubin (1992) method to the multivariate form. They propose the sth order moments other than the second order moment. This is calculated thus:

$$\hat{R}_s = \frac{\frac{1}{mn-1}\sum_{i=1}^{m}\sum_{t=1}^{n}|\bar{\theta}_i^t - \bar{\theta}|^s}{\frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{t=1}^{n}|\bar{\theta}_i^t - \bar{\theta}|^s};$$ (C.8)

for $s = 2, 3, 4, \ldots$. When $s = 2$ this is equal to the Gelman and Rubin (1992). They introduce an alternative to the $\hat{R}$ diagnostic by using interval lengths denoted by $\hat{R}_{interval}$ instead of the variance ratio. This is constructed as follows: from each individual chain, $100(1-\alpha)\%$ interval is taken, in other words, the $100\frac{\alpha}{2}\%$ and the $100(1-\frac{\alpha}{2})\%$ points of the n draws. This becomes the within-sequence interval length estimates, m. From the whole set of mn observations, secured from all the chains, calculate the $(100-\alpha)\%$ interval, which gives the total-sequence interval length estimate. Evaluate $\hat{R}$ defined as

$$\hat{R}_{interval} = \frac{length\ of\ total - sequence\ interval}{mean\ legth\ of\ the\ within - sequence\ intervals}$$ (C.9)

They state equation (D.9) is still the *Potential Scale Reduction Factor* except that it uses interval lengths instead of the variance ratio. It is also simpler than the original method.

They also provide the multivariate version of the variance-covariance matrix by:

$$\hat{V} = \frac{n-1}{n}W + (1 + \frac{1}{m})^{B}/_{n}$$ (C.10)

where

$$W = \frac{1}{m(n-1)}\sum_{j=1}^{m}\sum_{t=1}^{n}(\theta_j^t - \bar{\theta}_j)(\theta_j^t - \bar{\theta}_j)'$$

and

$$^{B}/_{n} = \frac{1}{m-1}\sum_{j=1}^{m}(\bar{\theta}_j - \bar{\theta})(\bar{\theta}_j - \bar{\theta})'$$

This denotes the $(q - dimensional)$ within and between-sequence covariance matrix estimates of the $q - variate$ functional θ, respectively. Brooks and Gelman (1998) state that convergence is achieved when the rotationally invariant distance between $\hat{V}$ and W are "sufficiently" close. They therefore introduce a scalar measure of this distance by using the maximum root statistics formula:

$$\hat{R}^p = \frac{n-1}{n} + (\frac{m+1}{m})\lambda_1$$ (C.11)

where λ_1 is the largest eigenvalue of the symmetric, positive definite matrix $W^{-1}B/n$.

The quantity $\hat{R}^p$ is the multivariate *Potential Scale Reduction Factor* (MPSRF) as against the univariate measure of Gelman and Rubin (1992).

C.2.2. Spectral Methods

The diagnostic begins by assuming that, in the steady state of a Markov chain, a covariance stationary process exists. The aim is to test the null hypothesis of stationarity. If we have a sequence $K^t : t = 1, ..., n$ from a covariance stationary process with unknown spectral density, $S(\omega)$. Then, for $n \geq 1$, we let

$$J_0 = 0, \ J_n = \sum_{t=1}^{n} K^t \text{ and } \bar{K} = \frac{1}{n} J_n$$

and

$$\hat{B}_n(s) = \frac{(J_{[n,s]} - [n,s]\bar{K})}{(n\hat{S}(0))^{\frac{1}{2}}}, \ 0 \leq s \leq 1$$

where $\hat{S}(0)$ is an estimate of the spectral density based on the second half of the chain, in order to prevent an initial transient that tends to be too large.

For large n, $\hat{B}_n = \hat{B}_n(s) : 0 \leq s \leq 1$ is distributed approximately as a Brownian Bridge. Several statistics such as; the Cramer-von Mises statistic, Kolmogorov-Smirnov statistic (see Darling 1957 for explanation on these two tests), and, the Schruben's statistic (Schruben et al. 1983) are used to test the null hypothesis of stationarity. They suggest the use of these statistics in estimating the length of the burn-in. This diagnostic is used for univariate observations and only single replications. However, Brooks and Gelman (1998) suggests a generalisation of the diagnostic by making use of the multivariate observations and multiple replications analogues of the Kolmogorov-Smirnov statistic using the methods stated in Conover (1965, 1967), and Ahmad (1976) for example.

To arrive at the functional form of g_i, we let $K(X', X)$ denote the probability of moving from X' to $X^{(i+1)} = \theta$ in one iteration of the Gibbs sampler. Then, if g_i is the joint density of the observations sampled at iteration i, then the joint density of the observations sampled at the next iteration (g_{i+1}) is given by

$$g_{i+1} = \int K(\boldsymbol{\theta'}, \boldsymbol{\theta}) g_i(\boldsymbol{X'}) d\lambda(\boldsymbol{X'}) \tag{C.12}$$

where $K(\boldsymbol{\theta'}, \boldsymbol{\theta})$ denotes a Gibbs transition density. One may approximate equation (D.12) using the Monte Carlo sum given as:

$$g_{i+1}(\theta) \approx \frac{1}{m} \sum_{j=1}^{m} K(\boldsymbol{\theta^j}, \boldsymbol{\theta}) \tag{C.13}$$

where $\theta^j : j = 1, 2, ..., m$. Ritter and Tanner (1992) state that θ^j may be a sample from m parallel chains at iteration $i + 1$ or from a batch of size m of consecutive iterates in a single chain. Ritter and Tanner (1992) suggest multiple draws (replication) of the diagnostics and computing certain statistics like the interquartile range or standard deviation of the weights at equally spaced intervals and then monitoring this value as the iteration increases. A histogram of the chosen statistic is plotted at different iterations until the fluctuate about a value and the weight histogram plot fluctuates about a "spike distribution".

If any two points within the parameter space is available say (α_1, β_1) and (α_2, β_2) we can write the "anchored ratio convergence criterion" as:

$$\hat{\theta}_A = \frac{\hat{p}(\alpha_1)p(\beta_1|\alpha_1)}{\hat{p}(\alpha_2)p(\beta_2|\alpha_2)}; \qquad \hat{\theta}_B = \frac{\hat{p}(\alpha_1)p(\beta_1|\alpha_1)}{\hat{p}(\alpha_2)p(\beta_2|\alpha_2)} \tag{C.14}$$

Zellner and Min state that if the values of the Gibbs sampler output is "satisfactory" then the values of $\hat{\theta}_A$ and $\hat{\theta}_B$ in equation (D.14) will be close to the value of θ obtained below:

$$\theta = \frac{\pi(\alpha_1, \beta_1)l(\alpha_1, \beta_1|y)}{\pi(\alpha_2, \beta_2)l(\alpha_2, \beta_2|y)} \tag{C.15}$$

where π and l are the prior density and likelihood respectively.

They state the "ratio convergence criterion" as when either $\hat{\theta}_A$ or $\hat{\theta}_B$ defined in equation (D.15) above equals one.

C.2.3. Normed Distance Criteria

This is created by the inversion of the Gibbs sampler (please see Robert and Casella 2004 p.478 for more exposition). Thus, his method found a way to prove that under certain regularity condition:

$$\|ft - f\| \overset{n \to \infty}{\Longrightarrow} 0 \tag{C.16}$$

Roberts showed that running m parallel Gibbs sampler chains with same starting values $\theta^{(0)}$, $\|f_t - f\| + 1$ is an unbiased estimator which is defined as:

$$J_t = \frac{1}{m(m-1)}\sum_{l \neq p} \frac{\underline{K}(\tilde{\theta}_l, \tilde{\theta}_p^{(2t-1)})}{f(\theta_p^{(2t-1)})} \tag{C.17}$$

where $t = 0, 1, ...,$, $\underline{K}$ is the transition kernel density for the "backward" half of the reversible sampler, l and p replications of the chains, $\tilde{\theta}_l$ is the forward half of the first sampler, $\tilde{\theta}_p^{(t)}$ is the value obtained after the tth draw of the pth chain. As stated in Robert and Casella (2004), because the distribution f is known up to the multiplicative constant, the limiting value of J_t is unknown, thus the convergence, J_t needs to be evaluated graphically. Roberts (1996), Brooks and Gelman (1998), and, Brooks and Roberts (1998) have modified equation (D.17) further by introducing additional diagnostics given as:

$$I_t = \sum_{i=1}^{m} \frac{\underline{K}(\tilde{\theta}_l, \tilde{\theta}_i^{(2t-1)})}{f(\theta_p^{(2t-1)})} \tag{C.18}$$

Roberts (1996) recommends that one monitors J_t and I_t until the point that they converge to the same point. Brooks and Roberts (1998) state that J_t is more interpretable if it is calculated as:

$$\frac{1}{m-1}\sum_{l \neq p} \frac{\underline{K}(\tilde{\theta}_l, \tilde{\theta}_p^{(2t-1)})}{f(\theta_p^{(2t-1)})} \tag{C.19}$$

If we let $x^t : t = 0, 1, 2, ...$ be a sequence of output from the d-dimensional the Markov chain sampler with $\pi(x)$ as its target (stationary) density with respect to the *Lebesgue measure*. The target density is defined as $\pi(x) = \theta g(x)$ with $g(x)$ known, non-negative and integrable and θ the inverse of the normalisation constant. Yu proposes the following procedures

First, given a one-dimensional bounded symmetric kernel $K(\cdot)$ such that $\int_{R^d} K(|x|)dx = 1$, let $h(\cdot)$ be the d–dimensional kernel based on K with $\sigma > 0$ and $|\cdot|$ being the Euclidean norm in R^d such that:

$$h_\sigma(x) = \frac{1}{\sigma^d}K(\frac{|x|}{\sigma}) \tag{C.20}$$

The kernel estimator of $\pi(\cdot)$ with bandwidth b_n is defined as

$$\hat{\pi}_n(x) = \frac{1}{n} \sum_{i=1}^{n} hb_n(\boldsymbol{x} - \boldsymbol{X_i}) \tag{C.21}$$

An estimate of the normalisation constant is given by:

$$\hat{\theta}_\sigma = \frac{1}{n(n-1)} \sum_{i \neq j} h_\sigma \frac{(X_i - X_j)}{g(X_j)} \tag{C.22}$$

A sub-set, A with non-zero Lebesgue measure of the support of the stationary distribution, π with increment time value, t_0. Starting at $t = t_0$ and estimating the bandwidth b_t based on the data by modifying Silverman (1986), the L^1 distance over A between the kernel density estimator $\hat{\pi}$ and π is given by:

$$\hat{I}_t(A) = \int_A |\hat{\pi}_t(x) - \hat{\theta}g(x)| dx \tag{C.23}$$

Yu (1995) states that the process is then carried out at different times $t = t_0, 2t_0, 3t_0, ...$ until a plot of the $\hat{I}_t(A)$ is below a certain threshold which Yu gave as 0.3.

Brooks et al. (1997) began by running m independent chains, each of which is divided into blocks of n_0 observations. They introduce a variable x_i^t which denotes the state of chain i at time t. The $d-$dimensional transition kernel is defined as:

$$\underline{K}(\boldsymbol{x_i^t}, \boldsymbol{x}) = \prod_{j=1}^{d} \underline{K}(x_j | x_l\, l < j, \quad x_{il}^t\ \ l > j) \tag{C.24}$$

where l is a block, and, i and j can be seen as chains. For the lth block of the ith chain, they define

$$\underline{K}_{i,l}(\boldsymbol{x}) = \sum_{t=(l-1)n_0+1}^{ln_0} \frac{\underline{K}(\boldsymbol{x_i^t}, \boldsymbol{x})}{n_0} \tag{C.25}$$

The quantity, $\underline{K}_{i,l}(\boldsymbol{x})$ is the Rao-Blackwellised estimate of the density of X in block l of chain i. The mean of the distance between $\underline{K}_{i,l}$ and $\underline{K}_{j,l}$ denoted by $\hat{\delta}(\underline{K}_{i.l}, \underline{K}_{j,l})$, also referred to as the between chain mean is given as:

$$B_l = \frac{1}{m(m-1)} \sum_{i=1}^{m} \sum_{j \neq i} \hat{r}_{ij}(l) \tag{C.26}$$

where $\hat{r}_{ij}(l)$ is the estimate of the rejection rate, $r_{ij}(l)$ and it is calculated in two ways via:

$$\min\left(1, \frac{\underline{K}_{i,l}(\boldsymbol{x_s})}{\underline{K}_{j,l}(\boldsymbol{x_s})}\right) OR \quad 1 - \min\left(1, \frac{\underline{K}_{il}(\boldsymbol{x})}{\underline{K}_{jl}(\boldsymbol{x})}\right) \tag{C.27}$$

The first one is a sample $\boldsymbol{x_s}$ from the density $\underline{K}_{il}$ with $s = 1, 2, ..., n$ while the second one is a sample x from the density $\underline{K}_{jl}$.

C.2.4. Regeneration and Coupling Methods

Johnson (1996) built on the principle of coupling time introducing this diagnostic. In his illustration of this diagnostic, he defined $U = u_{ij}$, $j = 1, ..., p$, $t = 1, ...$ be an array of independent uniform random variables and $\mathbf{x^t} = \{x_{(j)}^t\}$ denote the tth sampled vector in a series of draws from the Gibbs sampler. Also, they defined $F(\cdot|\cdot)$ as a generic conditional distribution function derived from $\pi(\cdot)$, and let $F^{-1}(\cdot|\cdot)$ be the corresponding quantile function. With all the notation defined, he describe the coupling algorithm for the Gibbs sampler as follows:

Step 0   Choose c starting vectors, x_i^0, $k = 1, ..., c$ from an initial distribution P_0 and set $t = 1$.

Step 1   For $j = 1, ..., p$, $k = 1, ...c$, and continuous $x_{(j)}^t$, equate $x_{ij}^k = F^{-1}(u_{(j)}^t|x_{i(1)}^t, ..., x_{i(j-1)}^t, x$

For discrete chains x_{ij} set x_{ij}^k to the minimum value satisfying

$x_{ij}^k \leq F(u_{(j)}^t|x_{(i+1)1}^t, ..., x_{(i+1)(j-1)}^t, x_{i(j+1)}^{t-1}, ..., x_{i(k)}^{t-1}$

Step 2   For discrete chains, if $x_i^t = x_l^t$ $i \neq l$, $i, l \in \{1, ..., c\}$ then all the paths have converged. For the continuous chains, convergence occurs when $|x_i^t = x_l^t| \leq \varepsilon \quad \forall \ i \neq l$, $i, l\{1, ..., c\}$ and $\varepsilon > 0$. If the chain have converged after the t iteration, then we stop, otherwise return to step 1 and increase t by 1.

C.2.5. The Eigenvalue Bounds Methods

As stated in Cowles and Carlin (1996), and, Brooks and Roberts (1998), suppose there exists m parallel sample chains with K burn-in period and n total run length and if Z_n is the sample frequency we can estimate ρ, a_2, and, λ_2 by defining the value $\hat{\theta} = (\hat{\rho}, \hat{a}_2, \hat{\lambda}_2)$ which minimises:

$$S(\rho, a_2, \lambda_2) = \sum_{n=K+1}^{N} (\bar{Z}_n - \rho - a_2 \lambda_2^n)^2 \tag{C.28}$$

Garren and Smith (2000) proved that a consistent estimator of θ exists as N tends to infinity. Cowles and Carlin (1996) then state that one needs to plot the sample values of $\hat{\theta}$ and with the coinciding 95% bound for $K = 1, 2, ...$ and looking for the point where the asymptotic bias disappears. Garren and Smith (2000) state that as n increases the estimates become unstable (with dramatic increase in standard error); they suggested using this as the choice of the proper amount for the burn-in, K. This method is described by Brooks and Roberts (1998) as semi-empirical, this assertion is supported by Garren and Smith (2000) who described their method as some sort of theoretical standard.

First, they calculate U_t for each t iteration and form $Z_t = I(U_t \leq u)$ where $I(\cdot)$ is the indicator function, and, Z_t is a binary process that is obtained from the Markov chain by marginalisation and truncation, but it is not itself a Markov chain. However, it becomes approximately a Markov chain when we form the sub-sequence $Z_t^{(k)}$ where $Z_t^{(k)} = Z_{1+(t-1)k}$. They did not provide a formal proof for this, but they claim it is intuitively possible. We believe that this first step is probably the most important part of this method.

Afterwards, they determine the number of burn-in, $M = mk$ to be discarded as:

$$m = m^* = \frac{\log\left(\frac{\varepsilon(\alpha+\beta)}{\max(\alpha,\beta)}\right)}{\log \lambda} \tag{C.29}$$

where $\lambda = (1 - \alpha - \beta)$ and α and β are given through the transition matrix for $Z_t^{(k)}$ given as $\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$ and thus, $M = m^* k$.

To determine N, assuming large t, $\bar{Z}_t^{(k)}$ is approximately normally distributed with mean q and variance $\frac{1}{n} \frac{\alpha\beta(2-\alpha-\beta)}{\alpha+\beta}$. Then, N is found at the point where $P[q - r \leq \bar{Z}_t^{(k)} \leq q + r] = p$ is met, if

$$n = n^* = \frac{\frac{\alpha\beta(2-\alpha-\beta)}{(\alpha+\beta)^3}}{\left\{\frac{r}{\Phi(\frac{1}{2}(1+p))}\right\}} \tag{C.30}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function and $N = kn^*$. The value of q and r are fixed by the researcher. Raftery and Lewis (1992) fixed q and r at 0.025 and 0.005 respectively.

To determine k, they compare the first-order Markov chain with the second-order Markov chain model, and, choose the smallest value of k for which the first-order model is preferred. They then use the Bayesian Information Criteria (BIC) ratios to choose k. They state that instead of the Bayesian Information Criteria, other non-Bayesian methods can be used.

C.2.6. Distance Methods

Suppose $x_1, ... x_n$ is a sample from some probability density f, the kernel estimate can be written as:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} \underline{K}(\frac{x - x_i}{h}) \tag{C.31}$$

where $\underline{K}$ is a kernel density of choice by the researcher (they use the normal kernel); h is the bandwidth (standard deviation). They propose the 'Silverman (1986) rule of thumb' in choosing the standard deviation. Therefore, the *Hellinger* distance between two densities, f and g becomes:

$$\hat{H}(f,g) = [\frac{1}{2}\int(\sqrt{\hat{f}}(x)-\sqrt{\hat{g}(x)^2 dx})]^{1/2} \approx [\frac{1}{2}\sum_{i=1}^{k}(\sqrt{\hat{f}}(x_l)-\sqrt{\hat{g}(x_l)^2(x_l - x_{l-1})})]^{1/2} \tag{C.32}$$

where $\hat{H}(f,g)$ is the *Hellinger* distance estimate of the kernel density estimate.

They state that convergence occurs at the point where the Hellinger distance between the two densities tends to zero. They proved this assertion and also proved that the kernel density estimate of $\hat{H}(f,g)$ is accurate.

They use Anderson-Darling and the Kolmogorov-Smirnov methods to prove whether there was a difference between the distributions of successive chains. They test their method

using several diagnostics - the single-chain similarity diagnostics, the parallel chain similarity diagnostics, and, the robustness diagnostics.

They claim that one major advantage of their method is that it is not computationally burdensome. They affirm that their method has been able to reveal problems that other diagnostics have been able to reveal. These problems include the robustness, and, similarity between the distribution of successive chains. They claim that the ability of a researcher to decide on what kernel to use is an advantage of their method. They also affirm that their method is not limited to Markov chain methods and could be extended to other methods.

C.2.7. The Batch/Stratification Methods

The method is based on the fact that there exists some degenerate continuous distribution function $J(\cdot)$ such that

$$J_N(x) \to J(x) \tag{C.33}$$

for all x as N tends to infinity. $J_N(x) = \Pr[\tau_N||T_N - \theta|| \le x]$ where τ_n, $n = 1, 2, ...$ is an increasing sequence that diverges to ∞ as $n \to \infty$ and T_N is a statistic of interest.

Let $(X_1, X_2, ..., X_N)$ be an observed time series data from $X_s, s = 1, 2, ...$ such that we can generate two sets of approximately independent random variables $(X_i, X_{i+1}, ..., X_{i+n})$, and $(X_{i+n+k}, X_{i+n+k+1}, ..., X_{i+2n+k})$ where i and n are positive integers and k is the total length of the chain

If $T_{i,b}$ is the statistic of interest from the block $(X_i, X_{i+1}, ..., X_{i+b-1})$, the distribution of the "subsample values" $T_{i,b}$, $i = 1, ..., B$ is given by

$$L_N(x) = \frac{1}{B} \cdot \sum_{i=1}^{B} 1\{\tau||T_{i,b} - T_N|| \le x\} \tag{C.34}$$

If X_s is strong mixing and asymptotically stationary and equation (D.34) holds, then L_N is a consistent estimator of the limit distribution $J(\cdot)$ as long as $b \to \infty$ as $N \to \infty$ and as $b/N \to 0$ and $\tau_b/\tau_N \to 0$. Consistent estimation of the quantiles of $J(\cdot)$ can be achieved through the quantiles of $L_N(\cdot)$ for any $t\epsilon(0, 1)$ in probability as $N \to \infty$ by

$$L_N^{-1}(t) \to J^{-1}(t) \tag{C.35}$$

where $L_N^{-1}(t) \equiv \inf\{x : L_N(x) \geq t\}$ and $J^{-1}(t) = \inf\{x : J(x) \geq t\}$ are t quantiles of $L_N(\cdot)$ and $J(\cdot)$ respectively.

They propose stopping the simulation when the range of $(1 - \alpha)100$ confidence region (with $\alpha = 0.05$) is smaller than $0.001||\bar{X}_N||$.

Assuming there exists a sample size N, with choice parameter $\{\theta = \theta_1, ..., \theta_N\}$ generated from a stationary distribution π. The sample is divided into batches of size K, each with size n such that $N = K \times N$. A natural estimator of the mean is:

$$E_1 = \frac{1}{N} \sum_{i=1}^{N} \theta_i = \frac{1}{K} \sum_{k=1}^{k} \bar{\theta}_{(k)} \qquad (C.36)$$

where $\bar{\theta}_{(k)}$ is the Kth batch mean. A second estimator E_2 is then constructed based on stratification via similar means. The natural estimates of the variances of the asymptotic distributions of E_1 and E_2 are given by:

$$\hat{V}_1 = \frac{1}{n} \mathbf{d}_1' \hat{\sum}_z \mathbf{d}_1 \qquad (C.37)$$

and

$$\hat{V}_2 = \frac{1}{n} \mathbf{d}_2' \hat{\sum}_z \mathbf{d}_2 \qquad (C.38)$$

where $\mathbf{d}_1$ and $\mathbf{d}_2$ are gradients of the batches and $\hat{\sum}_z$ is an estimator of the variance of a multivariate normal distribution, $\sum_z$. The application of the ergodic theorem implies that $\mathbf{d}_2 - \mathbf{d}_1 \to 0 \ \ as \ \ n \to \infty$.

Convergence and mixing occurs at the point where $n\hat{V}_1$ and $n\hat{V}_2$ are equal. They also develop methods of determining the amount of burn-in required for a particular Markov chain. They state that the samples of $\hat{V}_1$ are normally more stable than that of $\hat{V}_2$. To accurate results they advised that researchers use large values of n and K since their method is based on the asymptotic property.

Philippe and Robert Riemann Sums Method

Suppose $Y = (X, Z) \in \mathbb{R} \times \mathbb{R}^{p-1}$ is distributed with known density g with X and Z defined as $X : x = x_1, ..., x_n$ and $Z : z = z_1, ..., z_n$. Let $\mathbb{E}^f[h(X)]$ be the parameter of interest defined as:

$$\mathbb{E}[h(X)] = \int_{\mathbb{R}} h(x)f(x)dx$$

where h is an integrable function and f is a probability distribution.

The Rao-Blackwellised Riemann sum estimator is given as:

$$\delta_T^{\text{R/RB}} = \sum_{t=1}^{T-1} (x^{(t+1)} - x^{(t)})h(x^{(t)})\hat{f}(x)^{(t)} = T^{-1} \sum_{t=1}^{T-1} ((x^{(t+1)} - x^{(t)})h(x^{(t)})h(x)^{(t)}) \left(\sum_{k=1}^{T} \pi(x^{(t)}|z^{(k)}) \right) \tag{C.39}$$

When convergence occurs, $\hat{f}$ converges to f. They state that in case the stationary distribution is only known up to a constant equation (D.39) is given by:

$$\delta_T^{\text{R/RB}} = \frac{\sum\limits_{t=1}^{T-1} (x^{(t+1)} - x^{(t)})h(x^{(t)})\hat{f}(x)^{(t)} \left(\sum_{k=1}^{T} \pi(x^{(t)}|z^{(k)}) \right)}{\sum_{t=1}^{T-1}(x^{(t+1)} - x^{(t)})(\sum_{k=1}^{T}(x^{(t)}|z^{(k)}))} \tag{C.40}$$

where $t = 1, 2, ..., T$

They state that their method can be used as a *control variate* technique in convergence assessment by using alternative versions of the Riemann sums.

C.3. The Markov chain Monte Carlo Diagnostics Results

C.3.1. Nigeria

This section presents the result of the autocorrelation, and the Raftery-Lewis diagnostic tests on the frontier parameters for an initial run of one thousand draws as suggested by Gelman et al. (2003). The diagnostics show the number of times we have to run our code for us to achieve convergence and the total number of burn-in needed. We start by presenting the results for the Nigeria data.

Table C.1.: Autocorrelation within each Frontier Parameter Chain

S/N	Variables	Lag 1	Lag 5	Lag 10	Lag 50
1	age	0.138	0.039	0.055	0.068
2	pesticide	0.089	0.048	0.026	0.063
3	equipment_used	0.200	0.134	0.128	0.070
4	malaria	0.122	-0.048	-0.013	0.013
5	constants	0.101	-0.000	0.015	-0.055

From table 9.1, it is seen that our variables exhibit slight autocorrelation lags 1, 5, 10, and, 50, which implies that autocorrelation is not a major problem in our covariates.

Next, we present the results of the Raftery-Lewis diagnostic test.

Table C.2.: Raftery-Lewis Diagnostics for each Frontier Parameter Chain

S/N	Variables	Thin	Burn	Total (N)	(Nmin)	I-stat
1	age	1	2	969	937	1.034
2	pesticide	1	2	969	937	1.034
3	equipment_used	1	2	969	937	1.034
4	malaria	1	2	969	937	1.034
5	constants	1	2	969	937	1.034

From table 9.2, it is seen that the thinning value in column three is one, this confirms the earlier result in table 9.1, that the sequence of draws almost lacks autocorrelation. The amount of burn-in required is small, as a result, we not did include any burn-in in our analysis.

The fifth column (N) indicates the number of draws needed to achieve convergence. This value is less than the 1,000 draws we used, which means the values obtained in this run should be accurate. The sixth column (Nmin) is the total draws needed if the draws are from an identically and independently distributed chain. This value is still less than the 1,000 draws we used. The *i-statistic* is the ratio of the fifth and sixth column and as stated by Raftery et al. (1992), a convergence problem results, if this value exceeds 5. Our value is less than 5 and thus our sampler does not suffer from convergence problem.

C.3.2. Ethiopia

Table C.3.: Autocorrelation within each Parameter Chain

S/N	Variables	Lag 1	Lag 5	Lag 10	Lag 50
1	landarea	0.638	0.204	0.083	0.044
2	fertilizer	0.735	0.623	0.625	0.618
3	landtype	0.485	0.203	0.206	0.175
4	asset	0.378	0.179	0.164	0.149
5	credit	0.432	0.015	-0.007	-0.011
6	slope	0.639	0.447	0.439	0.414
7	intercropping	0.386	0.023	0.002	0.014
8	malaria	0.592	0.359	0.357	0.326
9	constant	0.441	0.091	0.08	0.068

From table 9.3, our estimates exhibit weak to medium autocorrelation. We may need to run it for longer. The next diagnostics will tell us.

Table C.4.: Raftery-Lewis Diagnostics for each Parameter Cha**in**

S/N	Variables	Thin	Burn	Total (N)	(Nmin)	I-stat
1	landarea	1	2	976	937	1.04
2	fertilizer	1	2	976	937	1.04
3	landtype	1	2	976	937	1.04
4	asset	1	2	976	937	1.04
5	credit	1	2	976	937	1.04
6	slope	1	2	976	937	1.04
7	intercropping	1	2	976	937	1.04
8	malaria	1	2	976	937	1.04
9		1	2	976	937	1.04

From table 9.4, convergence occurs when we run our model for slightly less than 1000 with burn-in of two. Our initial run was 1000, thus we could reliably say the draws converged, we ran it for further 10,000 draws with no burn-in.

C.3.3. Tanzania

Table C.5.: Autocorrelation within each Parameter Chain

S/N	Variables	Lag 1	Lag 5	Lag 10	Lag 50
1	area	0.75	0.65	0.64	0.63
2	labour	0.41	0.13	0.15	0.13
3	otherinputs	0.46	0.25	0.24	0.24
4	erosioncontrol	0.66	0.56	0.56	0.53
5	extension	0.14	0.00	0.03	0.00
6	inputcredit	0.19	0.03	0.01	0.01
7	age	0.56	0.42	0.41	0.41
8	malariaprevalence	0.66	0.49	0.49	0.47
9	constant	0.14	0.02	0.03	0.00

The autocorrelation value varies across our covariates from lags 1,5, 10, and, 50. Some variables show high correlation values while others show little or no autocorrelation at all. We would have to investigate this further when we use the Raftery-Lewis diagnostics presently as this might be a convergence problem.

Table C.6.: Raftery-Lewis Diagnostics for each Parameter Chain

S/N	Variables	Thin	Burn	Total (N)	(Nmin)	I-stat
1	area	2	8	2402	937	2.56
2	labour	2	8	2402	937	2.56
3	otherinputs	2	8	2402	937	2.56
4	erosioncontrol	2	8	2402	937	2.56
5	extension	2	8	2402	937	2.56
6	inputcredit	2	8	2402	937	2.56
7	age	2	8	2402	937	2.56
8	malariaprevalence	2	8	2402	937	2.56
9	constant	2	8	2402	937	2.56

From table 9.6, it could be seen that 1000 draws is not enough to achieve convergence, we would need 2402 draws. The number of number of burn-in is eight while the *I-statistics* is less than 5 which is fine by Raftery-Lewis standard. The code we used for our analysis was run 10,000 times with no burn-in as the burn-in value is still negligible.

In the last few sections, we have managed to carry out a few pre-modelling exercises in order achieve reliable estimates, next, we present our results.

D. Data Collection And Description

D.1. The Data

D.1.1. The Nigeria Data Set

Sample Frame for the LSMS-ISA Survey

We present the distribution of the selection in table E.1 and E.2 below:

Table D.1.: Distribution of the Enumeration Areas

S/N	Zone	State	No of EAs Allocated	No of HOUSE HOLDs Selected	No of HOUSE-HOLDs Selected
1	North-Central	Plateau	11	110	110
		Kwara	12	120	120
		Niger	18	180	180
		Kogi	12	120	120
		benue	16	160	160
		Nassarawa	7	70	70
		FCT Abuja	4	40	40
		Sub-Total	80	800	800
2	North East	Borno	21	210	210
		Yobe	13	130	130
		Bauchi	17	170	170
		Gombe	8	80	80
		Taraba	9	90	90
		Adamawa	12	120	120
		Sub-Total	80	800	800
3	North West	Kaduna	12	120	120
		Jigawa	13	170	170
		Katsina	18	180	180
		Kano	20	200	200
		Kebbi	10	100	100
		Sokoto	8	80	80
		Zamfara	9	90	90
		Sub-Total	90	900	900
4	South East	Enugu	14	140	140
		Anambra	22	220	220
		Ebonyi	14	140	140
		Abia	11	110	110
		Imo	19	190	190
		Sub-Total	80	800	800
5	South-South	Cross-River	13	130	130
		Akwa-Ibom	15	150	150
		Rivers	21	210	210
		Bayelsa	7	70	70
		Edo	10	100	100
		Delta	14	140	140
		Sub-Total	80	800	800
6	South West	Oyo	23	230	230
		Osun	18	180	180
		Ogun	11	110	110
		Lagos	17	170	170
		Ekiti	8	80	80
		Ondo	13	130	130
		Sub-Total	90	900	900
		Total	500	5000	5000

*Adapted and modified from Nigerian NBS (2011) Information Document

Table D.2.: Distribution of Final Sample of 500 EAs and 5,000 Households for Panel Survey by State, Urban and Rural Sectors Within Each Zone

Zone	State	Total		Urban		Rural	
		No. of EAs	No. of HHs	No. of EAs	No. of HHs	No. of EAs	No. of HHs
North-Central	Benue	16	160	2	20	14	140
	Kogi	12	120	4	40	8	80
	Kwara	12	120	6	60	6	60
	Nasarawa	7	70	1	10	6	60
	Niger	18	180	4	40	14	140
	Plateau	11	110	2	20	9	90
	FCT Abuja	4	40	3	30	1	10
North-East	Adamawa	12	120	1	10	11	110
	Bauchi	17	170	3	30	14	140
	Borno	21	210	5	50	16	160
	Gombe	8	80	1	10	7	70
	Taraba	9	90	0	0	9	90
	Yobe	13	130	3	30	10	100
North-West Zone	Jigawa	13	130	2	20	11	110
	Kaduna	12	120	4	40	8	80
	Kano	20	200	3	30	17	170
	Katsina	18	180	3	30	15	150
	Kebbi	10	100	1	10	9	90
	Sokoto	8	80	2	20	6	60
	Zamfara	9	90	2	20	7	70
South-East	Abia	11	110	4	40	7	70
	Anambra	22	220	12	120	10	100
	Ebonyi	14	140	1	10	13	130
	Enugu	14	140	3	30	11	110
	Imo	19	190	2	20	17	170
South-South	Akwa-Ibom	15	150	4	40	11	110
	Bayelsa	7	70	1	10	6	60
	Cross-River	13	130	3	30	10	100
	Delta	14	140	4	40	10	100
	Edo	10	100	5	50	5	50
	Rivers	21	210	8	80	13	130
South-West	Ekiti	8	80	6	60	2	20
	Lagos	17	170	16	160	1	10
	Ogun	11	110	7	70	4	40
	Ondo	13	130	6	60	7	70
	Osun	18	180	14	140	4	40
	Oyo	23	230	15	150	8	80

*Culled from Nigerian NBS (2011 p. 18) Basic Information document

The Data Set File

Table D.3.: Household Data set Files

Section	Section Name	Dataset Filename
Cover	Cover	SECTA
1	Roster	SECT1
2	Education	SECT2
3	Labour	SECT3
4	Credit and Savings	SECT4
5	Household Assets	SECT5
6	Non-farm Enterprises	SECT6
7A	Meals Away From Home	SECT7A
7B	Household Food Expenditures	SECT7B
8	Household Non-food Expenditures	SECT81 – Expenditure in the past 7 days
		SECT82 – Expenditure in the past month
		SECT83 – Expenditure in the past six months
		SECT84 – Expenditure in the past twelve months
		SECT85 – Expenditure on twelve months recall
9	Food Security	SECT9
10	Other Income	SECT10

Culled from NNBS (2011, p.27)

Table D.4.: Community File

Section	Section Name	Dataset Filename
Cover	Cover	Contains Administrative data just like the household data file
1	Respondent Characteristics	SECTC1
2	Food Prices	SECTC2
3	Labour	SECTC3A: do individuals hire labour in the village
		SECTC3B: types of labour and amount paid to labour
4	Land Prices and Credit	SECTC4A: Land Prices
		SECTC4B: Credit

Culled from NNBS(2011, p.27)

<div align="center">

Table D.5.: Agriculture Data Set File

</div>

Section	Section Name	Dataset Filename
11a	Plot Roster	SECT11A - Ownership and Cultivation of land
		SECT11A1 – Plot details
11b	Land Inventory	SECT11B
11c	Input Costs	SECT11C
11d	fertiliser Acquisition	SECT11D
11e	Seed Acquisition	SECT11E
11f	Planting Field Crops	SECT11F
11g	Planting Tree Crops	SECT11G
11h	Marketing of Agricultural Surplus	SECT11H
11i	Animal Holdings	SECT11I
11j	Animal Costs	SECT11J
11k	Agriculture by-products	SECT11K
11l	Extension	SECT11L1 – Advice on topic and source
		SECT11L2 – Detail information of who in the household received the advice through the source
12	Network Roster	SECT12

*Adapted from Nigerian NBS (2011, p.28) Basic Information document

D.1.2. The Ethiopia Data Set

Selection of Sample for the Panel Survey

We reproduce the sampling frame, the actual percentage of samples selected from each farming system, the characteristics of each of the sample locations, the different times the survey was carried out in each Peasant Association in the Ethiopia Household survey from Dercon and Hoddinott (2011) below in tables E.6, E.7, and, E.8:

Table D.6.: The Sampling Frame of the Ethiopia Rural Household Survey

Farming Systems in Ethiopia	Population Share in 1994 (%)	Sampling Share in 1994 (%)	Sampling Share in 1989	Number Villages in 1989	Number of Villages in 1989 and 1994	Panel House-hold (%)
Grain-Plough Complex - Northern Highlands	21.2	20.2	-	3	0	-
Grain-Plough Complex - Central Highlands	27.7	29.0	31.0	4	2	32.4
Grain Plough Arsi/Bale	9.3	14.3	25.4	2	1	25.6
Sorghum Plough/Hoe Hararghe	9.9	6.6	15.0	1	1	12.4
ensette *(with or without coffee/cereal)*	31.9	29.9	8.7	5	2	29.6
Total	100 %	100%	100%	15	6	100%

Table D.7.: Attributes of the Survey Locations

Survey Site	Location	Background	Main Crops
Haresaw	Tigray	Poor and vulnerable area	Cereals
Geblen	Tigray	Poor and vulnerable area	Cereals
Dinki	North Shewa	Badly affected in 1984/85 famine	Millet, Teff
Debre Berhan	North Shewa	Highland site; near town	Teff, Barley, Beans
Yetmen	Gojjam	Ox-plough cereal farming system of highlands	Teff, Wheat, and Beans
Shumsha	S.Wollo	Poor area in neighbourhood of airport near Lalibela	Cereals
Sirbana Godeti	Shewa	Rich area, target of agricultural policy. Cereal, ox-plough system	Teff
Adele Keke	Hararghe	Highland site, experiences drought in 1985/86	Millet, Maize, Coffee, Chat
Korodegaga	Arssi	Poor cropping area in neighbourhood of rich valley	Cereals
Turfe Kechemane	South Shewa	Highlands. Ox-plough, rich cereal area	Wheat, barley, Teff, Potato
Imdibir	Shewa (Gurage)	Densely populated ensette area	ensette, Chat, Coffee, Maize
Aze Deboa	Shewa (Kembata)	Densely populated with long tradition of substantial seasonal and temporary migration	ensette, coffee, maize, sorghum
Addado	Sidmano (Dilla)	Rich coffee producing area and densely populated	Coffee, ensette
Gara Godo	Sidmano (Wolayta)	Densely packed ensette-farming area; experiences famine in 1983/84; malaria in mid-88	Barley, ensette
Doma	Gama Gofa	Resettlement area (1985); semi-arid, and experiences droughts in 1985,88,89,90; and remote	ensette, Maize

Dercon et al.(2011) cites Webb and von Braun, (1994); Bevan and Pankhurst, (1996)

Table D.8.: Timing of Activities for the Survey

Survey site	Location	Main Harvest Time of Interview	Rnd 1 1994	Rnd 2 1994-95	Rnd 3 1995	Rnd 4, 1997	Rnd 5, 1999	Rnd 6, 2004
Haresaw	Tigray	Oct. -Nov.	June-July	Jan.	Mar.	June	August	April
Geblen	Tigray	Oct. -Nov.	June-July	Jan.	Mar.	June	Sept.	May
Dinki	N.Shoa	Dec.	Mar. - April	Nov.	Jan.	Oct, Nov	August	May, June
Debre berhan	N.Shoa	Nov. - Dec.	Mar. - April	Oct.	Mar.	June-August	July	April, May
Yetmen	Gojjam	Nov. - Dec.	Mar. - April	Oct.	Mar.	Sept., Oct.	June	may, June
Shumsha	S.Wollo	Oct. - Dec.	June - July	Dec. - Jan.	May	Oct., Nov.	Sept.	April, May
Sirbana Godeti	Shoa	Nov. - Dec.	Mar - April	Nov.	Mar.	June, July	June	April, May
Adele Keke	Hararghe	Nov. - Dec.	May-June	Oct.	April	Oct., Nov.	June	April, May
Koro-degaga	Arssi	Oct. - Nov	May-June	Nov. - Dec.	May - June	June, July	August	May
Turfe Keche-mane	S.Shoa	Dec.	Mar. - April	Sept. - Oct.	March-April	Sept., Oct.	June	May, June
Imdibir	Shoa(Gurage)	Oct. - Dec.	Mar - April	Oct.	March	June, July	June	April, May
Aze Deboa	Shoa(Kembata)	Oct. - Nov.	Mar - April	Sept. - Oct.	March	Sept., Oct.	June	July
Addado	Sidamo (Dilla)	Dec. - Jan.	Mar.-April	January	March	June, July	Sept.	April, May
Gara Godo	Sidamo (Wolayta)	Aug. - Dec.	Mar. - May	Oct.	March	June, July	June	April, May
Doma	Gama Gofa	Sept. - Dec.	May-June	Dec.	May-June	Nov.	Sept.	May, June

Source: Dercon et al.(2011) cites Community survey ERHS and Bevan and Pankhurst (1996)

Merging Each of the Data Files

We present a sample of the survey data files for round 7 below:

Table D.9.: Prototype of the Data Files For The Ethiopia Rural Household Survey

Data File Name	Data file link to questionnaire part and section	File description	Additional information, if any
PART 1. HOUSEHOLD DEMOGRAPHICS, HOUSEHOLD ASSETS AND NON- AGRICULTURAL INCOME			
R7p1_s0.dta	Part 1, cover page	Respondent, interview date information for part 1, section 1	
Roster_r7.dta	Roster file	Roster file	
R7p1_s1a.dta	Part 1, section 1A	Household demographics, continuing members for households with roster card	
R7p1_s1a_YYrv2.dta	Part 1, section 1A	Updated R7p1_s1a.dta	Updated the file after accessing data questionnaire
R7p1_s1b.dta	Part 1, section 1B	Updating household demographics, new members	
R7p1_s1b_YYrv2.dta	Part 1, section 1B	Updated R7p1_s1b.dta	Updated the file after accessing data questionnaire
R7p1_s1c.dta	Part 1, section 1C	Updating household demographics, former members for households with roster card	
R7p1_s1c_YYrv.dta	Part 1, section 1C	Updated R7p1_s1c.dta	Updated the file after accessing data questionnaire
R7p1_s1d1.dta	Part 1, section 1D	Updating household demographics, child mortality	Question1

R7p1_s1d2.dta	Part 1, section 1D	Updating household demographics, child mortality	Questions 2-5
Part 1, section 2: Children's activities and education			
R7p1_s2a.dta	Part 1, section 2	Children's activities and education.	Questions 1-7
R7p1_s2b.dta	Part 1, section 2	Children's activities and education.	
Questions 8-10	Household level		
Part 1, section 3: Assets			
R7p1_s3a.dta	Part 1, section 3	Assets, Question 1-3	
R7p1_s3b.dta	Part 1, section 3	Assets, Question 4	
R7p1_s3c.dta	Part 1, section 3	Assets, Question 4a , 4c	Columns 4b and 4c assigned to entries next to question 4a
R7p1_s3d.dta	Part 1, section 3	Assets, Question 5a-5c	
R7p1_s3e.dta	Part 1, section 3	Assets, Question 6	
R7p1_s3f.dta	Part 1, section 3	Assets, Question 7-10	
Part 1, section 4: Credit			
R7p1_s4a.dta	Part 1, section 4	Credit, Questions 1-2	
R7p1_s4b1.dta	Part 1, section 4	Credit, Questions 3-11	
R7p1_s4b2.dta	Part 1, section 4	Credit, Questions 12-20	
R7p1_s4c.dta	Part 1, section 4	Credit, Question 21	
R7p1_s4d1.dta	Part 1, section 4	Credit, Questions 22-28	
R7p1_s4d2.dta	Part 1, section 4	Credit, Questions29-35	
R7p1_s4e.dta	Part 1, section 4	Credit, Questions 36-37	

	Part 1, Section 5: Non-food expenditures		
R7p1_s5.dta	Part 1, section 5	Non-food expenditures, Questions 1-3	
	Part 1, Section 6: Off-farm income and business activities		
R7p1_s6a.dta	Part 1, section 6	Questions 1a, 1b	
R7p1_s6b.dta	Part 1, section 6	Questions 2-7	
R7p1_s6c.dta	Part 1, section 6	Questions 1	
R7p1_s6d.dta	Part 1, section 6	Questions 2-13	
R7p1_s6e.dta	Part 1, section 6	Questions 14a, 14b	
R7p1_s6f.dta	Part 1, section 6	Questions 15-18	
R7p1_s6g.dta	Part 1, section 6	Questions 19	
R7p1_s6h1.dta	Part 1, section 6	Question 20-28	
R7p1_s6h2.dta	Part 1, section 6	Questions 29-38	
R7p1_s6i.dta	Part 1, section 6	Questions 39	
R7p1_s6j1.dta	Part 1, section 6	Questions 40-47	
R7p1_s6j2.dta	Part 1, section 6	Questions 48-57	
	Part 1, Section 7: Men's perception of poverty and well-being		
R7p1_s7a.dta	Part 1, section 7	Questions 1-21	
R7p1_s7b.dta	Part 1, section 7	Questions 22-38	
R7p1_s7c.dta	Part 1, section 7	Questions 39a-41	
	PART 2: AGRICULTURE		
R7p2_resp.dta	Part 2, cover page	Respondent information and interview dates of agricultural section	
	Part 2, section 1A: Land and its use - quality of land and crops grown		
R7p2_s1a.dta	Part 2, section 1A		
	Part 2, section 1B: Land and its use - Land acquisition and rights		
R7p2_s1b1.dta	Part 2, section 1B	Questions 1-13	
R7p2_s1b2.dta	Part 2, section 1B	Questions 14-15	
	Part 2, section 1C:Use of Inputs		
R7p2_s1c.dta	Part 2, section 1C		
	Part 2, section 1D: Land and its use- Plot outputs and sales		
R7p2_s1d1.dta	Part 2, section 1D	Question 1-5	
R7p2_s1d2.dta	Part 2, section 1D	Questions 6	
R7p2_s1d3.dta	Part 1, section 6		
R7P2_s1e1.dta	Part2, section 1E	Question 1	
R7P2_s1e2.dta	Part2, section 1E	Questions 2-6	
	Part 2, section 2A: Agricultural inputs-labor sharing		
R7p2_s2a1.dta	Part 2, section 2A	Questions 1-2	
R7p2_s2a2a.dta			
R7p2_s2a2b.dta	Part 2, section 2A	Questions 15-25	
	Part 2, section 2B: Agricultural inputs-family and hired labor		
R7p2_s2b1.dta	Part 2, section 2B	Questions 1-22	
R7p2_s2b2.dta	Part 2, section 2B	Questions 23-26	
	Part 2, section 2C: Agricultural inputs-Other expenditures		
R7p3_s8b.dta	Part 3, section 8	Question 19-36	
R7p3_s8c.dta	Part 3, section 8	Questions 37-39	
	Part 3, section 9: Participation in village life and decision making		
R7P3_s9.dta	Part 3, section9		

Part 3, section 10: Participation in household decision making			
R7P3_s10a.dta	Part3, section109	Question 1	
R7P3_s10b.dta	Part3, section 10	Questions 2-4	
R7P3_s10c.dta	Part3, section 10	Question 5	
R7P3_s10d.dta	Part3, section 10	Question 6-7	
R7P3_s10e.dta	Part3, section 10	Question 8-24	
Part 3, section 11: Allocation of Assets on Divorce			
R7P3_s11a.dta	Part3, section 11	Questions 1-7	
R7P3_s11b.dta	Part3, section 11	Questions 8	
Part 3, section 12: Knowledge of Land Rights resulting from Land Registration			
R7P3_s12.dta			
PART 4: SHOCKS, PUBLIC WORKS, DROUGHT, NETWORKS, IDDIR AND TRUST			
R7P4_resp.dta	Part 4 cover page	Respondent and interview dates information	
Section 1: Shocks and Coping Mechanisms			
R7p4_s1a.dta	Part4 section 1: Long term shocks and coping mechanism	Questions 1, shock codes 101-601	
R7p4_s1b.dta	Part4 section 1: Long term shocks and coping mechanism	Last question on Shocks, "Thinking about the last ten years......"	
Part4, section 2: Access to the Productive Safety Nets Program, Public works			
R7p4_s2a.dta	Part4 section 2	Questions 1, 2	
R7p4_s2b.dta	Part4 section 2	Questions 3	
R7p4_s2c.dta	Part4 section 2	Questions 4	
Part4, section 3: Networks			
R7p4_s3a.dta	Part4 section 3	Questions 1-22	
R7p4_s3b.dta	Part4 section 3	Questions 23-24	
Section 4, section 4: IDDIR			
R7p4_s4a.dta	Part4 section 4	Questions 1-3	
R7p4_s4b.dta	Part4 section 4	R7p4_s4c.dta	Part4 section 4

D.1.3. The Tanzania Data

Because we had discussed the LSMS-ISA data in the Nigeria part, we immediately present the data files and the sampling design of the Tanzania version of the survey below:

The Data File

The data files for the survey are divided into four parts according to the number of different types of questionnaire administered. Hence, we have the household, agriculture, community, and, the fisheries data files. In each of the data files the prefix *SEC* was used and in the case of the household and the community data files an alphabet follows, while, for the agriculture data files a number and then an alphabet follows. We show examples of this in tables E.10, E.11 and E.12 below:

271

Table D.10.: Tanzania Household Questionnaire

Household Questionnaire		
Date File	Description	Unique Identifier
SEC_A_T.dta	Household Identification (Section A), GPS (Section T)	hhid
SEC_B_C_D_E1_F_G1_U.dta	Household Roster (Section B), Education (Section C), Health (Section D), Labour (Section E), Food Consumption Outside Household (Section F), Anthropometrics (Section U)	hhid, sbmemno
S(continued on next page)EC_E2.dta	Additional labour enterprises (Sub section of E, q24-44)	hhid, namba_ya_biashara
SEC_G2.dta	Children Living Elsewhere (Section G)	hhid, sgchildno
SEC_H1_J_K2_O2_P1_Q1_S1.dta	Housing, Water and Sanitation (Section J), Governance (Section H, q1-3, 10,11), Food Consumption (Section K, q7-9), Assistance and Groups (Section O, q6), Partial Crime and Justice (Section Q, q1-3, q16-20)	hhid
SEC_H2.dta	Governance (Section H, q4-7)	hhid
SEC_H3.dta	Governance (Section H, q8, 9)	hhid, shmeet
SEC_I.dta	Violence Against Women (Section I)	hhid, siq1
SEC_K1.dta	Food Consumption (Section K)	hhid, skcode
SEC_L.dta	Non-Food Expenditures (Section L)	hhid, slcode
SEC_M.dta	Non-Food Expenditures (Section M)	hhid, smcode
SEC_N.dta	Household Assets (Section N)	hhid, sncode
SEC_O1.dta	Assistance and Groups (Section O, q1-5)	hhid, socode
SEC_O3.dta	Assistance and Groups (Section O, q7-15)	hhid, somemno
SEC_P2.dta	Credit (Section P)	hhid, sp2code
SEC_Q2.dta	Crime and Justice (Section Q, q4-25)	hhid, sq2code
SEC_Q3.dta	Crime and Justice (Section Q, q21-28)	hhid, sqq21
SEC_R.dta	Recent Shocks (Section R)	hhid, srcode
SEC_S2.dta	Deaths in the Household (Section S)	hhid, sscode

Table D.11.: Agriculture Questionnaire

Date File	Description	Unique identifier
SEC_1_ALL.dta	Household Roster (Section 1)	hhid, rosterid
SEC_2A.dta	Plot Roster (Section 2A)	hhid, plotnum
SEC_2B.dta	Plot Roster (Section 2B)	hhid, plotnum
SEC_3A.dta	Plot Details (Section 3A)	hhid, plotnum
SEC_3B.dta	Plot Details (Section 3B)	hhid, plotnum
SEC_4A.dta	Crops by Plot (Section 4A)	hhid, plotnum, zaocode
SEC_4B.dta	Crops by Plot (Section 4B)	hhid, plotnum, zaocode
SEC_5A.dta	Crops – Household Totals (Section 5A)	hhid, zaocode
SEC_5B.dta	Crops – Household Totals (Section 5B)	hhid, zaocode
SEC_6A.dta	Permanent Crops by Plot (Section 6A)	hhid, plotnum, zaocode
SEC_6B.dta	Permanent Crops by Plot (Section 6B)	hhid, plotnum, zaocode
SEC_7A.dta	Permanent Crops by Crop (Section 7A)	hhid, zaocode
SEC_7B.dta	Permanent Crops by Crop (Section 7B)	hhid, zaocode
SEC_8A.dta	Outgrower Scheme and Contract Farming (Section 8A)	hhid, zaocode
SEC_8B.dta	Outgrower Scheme and Contract Farming (Section 8B)	hhid, zaocode
SEC_8C.dta	Outgrower Scheme and Contract Farming (Section 8C)	hhid, zaocode
SEC_9_ALL.dta	Processed Agricultural Products and Agricultural Bi-Products (Section 9)	hhid, zaocode, byproduct
SEC_10A.dta	Livestock (Section 10A)	hhid, animal

*Adapted from The Tanzania NBS(2008) Information document

SEC_10B.dta	Livestock By-Products (Section 10B)	hhid, lvstkprod
SEC_11_ALL.dta	Farm Implements and Machinery (Section 11)	hhid, impcode
SEC_12A.dta	Partial Fishery and Aquaculture (Section 12, q2-9)	hhid, itemcode
SEC_12B.dta	Partial Fishery and Aquaculture (Section 12, q10-15)	hhid, itemcode
SEC_12C.dta	Partial Fishery and Aquaculture (Section 12, q24-26)	hhid, areacode
SEC_12D.dta	Fishery and Aquaculture (Section 12, q16-20)	hhid, itemcode
SEC_13A.dta	Extension (Section 13, q1-6)	hhid, source
SEC_13B.dta	Extension (Section 13, q7-9)	hhid, source
SEC_QNFLOW.dta	Household Identification (Section 1A), Plot Roster (Section 2, q1, 5, 6), Plot Details (Section 3, q1), Permanent Crop by Plot (Section 7, q1), Outgrower (Section 8, q1), Processed Agricultural Products (Section 9, q1), Livestock (Section 10A, q1), Farm Implements (Section 11, q1), Fishery and Aquaculture (Section 12, q21-23), Extension (Section 13, q11-15)	hhid
SEC_PLOTGPS.dta	GPS Data	hhid
SEC_NETWORK.dta	Network Roster Card	hhid, nid

Table D.12.: Community Questionnaire

Date File	Description	Unique Identifier
SECTA1A2.dta	Community Identification (Section A)	Community
SECTCB.dta	Access to Basic Services (Section B)	Service
SECTCC.dta	Investment Projects (Section C)	Investment Project
SECTCD.dta	Land Use (Section D)	Community
SECTCEFG.dta	Agriculture (Section E), Demography and Family Issues (Section F), Governance (Section G)	Community
SECTCH.dta	Roster of Community Leaders (Section H)	Community Leader
SECTCI.dta	Crime and Policing (Section I)	Community
SECTCJ.dta	Market Prices (Section J)	Item

E. MATLAB© Exercises