

**ANALYSIS OF SPATIAL DIFFERENTIALS OF INCOME INEQUALITY IN
GHANA: APPLICATION OF BAYESIAN ESTIMATION**

BY

BENEDICTA TUTUANI

(10227964)

**THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA,
LEGON IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE
AWARD OF MPhil STATISTICS DEGREE**

JULY, 2015

DECLARATION

Candidate's Declaration

This is to certify that, this thesis is the result of my own work and that no part of it has been presented for another degree in this University or elsewhere.

SIGNATURE:

DATE:

BENEDICTA TUTUANI

(10227964)

Supervisors' Declaration

We hereby certify that this thesis was prepared from the candidate's own work and supervised in accordance with guidelines on supervision of thesis laid down by the University of Ghana.

SIGNATURE:

DATE:

DR FELIX OKOE METTLE

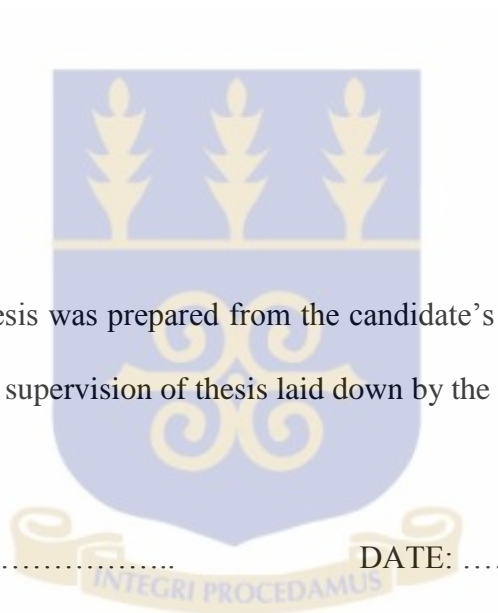
(Principal Supervisor)

SIGNATURE:

DATE:

DR SAMUEL IDDI

(Co-Supervisor)



ABSTRACT

The prime objective of the study was to introduce the application of a Bayesian method of estimation in the computation of the Gini coefficient. The bootstrap estimation technique was also employed to obtain confidence intervals for the estimated Gini coefficients which were used to statistically analyze the significant change of inter-household income inequality between 2005/06 and 2012/13 in Ghana with emphasis on the sex of head of household. Most of the methods used in calculating the Gini coefficient are numerically determined and many rely on income. This study provides a more statistical way of calculating the Gini coefficient which depends on income and household size. The income gap between the rich and the poor in Ghana increased from 0.507 in 2005/06 to 0.647 in 2012/13. Meanwhile, the percentage change in the unequal distribution of income was higher among all male headed households than their female counterparts. It was also discovered that not only was the spatial differentials of inter-household income inequality between 2005/06 and 2012/13 alarming but also statistically significant. Income inequality remains a threat to the developmental strategies in Ghana. Government and policy makers need to formulate policies and programs targeted at bridging the income gap between the rich and the poor in Ghana.

DEDICATION

I dedicate this work to God Almighty.



ACKNOWLEDGEMENT

I thank the almighty God for how far He has taken me in academia. Had it not been His grace and mercies nothing would have been possible.

Special thanks to UG-Carnegie Next Generation of Academics in Africa Project for their support and contribution towards the successful completion of this thesis. Through the Carnegie scholarship grant, I was able to purchase a laptop with the required specifications needed to run my analysis without which the study will not have been fruitful.

I would like to acknowledge the relentless effort of Dr. Felix O. Mettle in assisting in putting this document together.

I would also like to express my heart felt gratitude to Dr. Samuel Iddi, Dr. Ezekiel N. N. Nortey, Dr. K. Doku-Amponsah and Mr. Enoch N. B. Quaye for their assistance and encouragement.

I am deeply indebted to my godfather Mr. David Y. Mensah for extending a helping hand whenever the need arises. To all my friends especially to my fellow Carnegie scholars, I am forever grateful.

God bless you all.

TABLE OF CONTENT

	Page
Table of contents	
DECLARATION	i
ABSTRACT	ii
DEDICATION	iii
ACKNOWLEDGEMENT	iv
LIST OF FIGURES	ix
LIST OF TABLES	x

CHAPTER ONE: INTRODUCTION

1.0 Background of study.....	1
1.1 Statement of the problem	3
1.2 Objectives of the study.....	5
1.3 Significance of the study	6
1.4 The organizational structure	6

CHAPTER TWO: LITERATURE REVIEW

2.0 Introduction	7
2.1 Definition of inequality?	7
2.2 Measures of income inequality	8
2.3 The Gini coefficient	8
2.4 Properties of the Gini coefficient	9
2.5 Measures of Gini coefficient	10
2.6 Theoretical framework	12
2.7 The decomposition of income inequality	15

2.8 Empirical results	17
2.9 Definition of poverty?	18
2.10 The relationship between income inequality and poverty.....	19
2.11 Factors that influence income inequality	20
2.12 Significance of Spatial analysis	23

CHAPTER THREE: METHODOLOGY

3.0 Introduction	25
3.1 Source of data.....	25
3.2 Variables	26
3.3 Data organization.....	26
3.4 Derivation of the Lorenz curve	27
3.5 The regression model	29
3.5.1 Assumptions of the regression model	30
3.6 Bayesian estimation	30
3.7 Estimating μ and Σ of the prior distribution	37
3.8 Model estimation	38
3.9 Model evaluation	39
3.10 Bootstrapping the Gini coefficient	40
3.11 Significance test for differences in the Gini coefficient	41

CHAPTER FOUR: ANALYSIS OF DATA AND DISCUSSION OF RESULTS

4.0 Introduction	42
4.1 Data processing	42
4.2 Descriptive analysis of data	42

4.2.1 Sample sizes.....	43
4.2.2 Descriptive statistics of income and household size.....	43
4.2.2.1 Income.....	44
4.2.2.2 Household size	45
4.3 Estimation of Gini coefficient	46
4.3.1 The Prior distribution.....	46
4.3.2 The Posterior distribution	47
4.4 Income inequality in Ghana during 2005/06	51
4.4.1 Regions	51
4.4.2 Ecological zones	52
4.4.3 Urban/ Rural localities	53
4.5 Income inequality in Ghana during 2012/13	54
4.5.1 Regions.....	54
4.5.2 Ecological zones	55
4.5.3 Urban/ Rural localities	56
4.6 Summary statistic of income inequality for 2005/06 and 2012/13 in Ghana	57
4.7 Spatial differentials of income inequality.....	58

CHAPTER 5: SUMMARY, CONCLUSION AND RECOMMENDATION

5.0 Introduction	63
5.1 Summary	63
5.2 Conclusion	64
5.3 Recommendation	66

References	67
Appendix 1	75
Appendix 2	77
Appendix 3.....	83

LIST OF FIGURES

Figure 1: The Lorenz curve.....	13
Figure 2: Graphical representation of the fitted Lorenz curve	48
Figure 3: Graphical representation from bootstrap sample	50

LIST OF TABLES

Table 1: Summary statistic of household income.....	44
Table 2: Summary statistic of household size	45
Table 3: Gini coefficient by regions in 2005/06.....	51
Table 4: Gini coefficient by ecological zones in 2005/06	52
Table 5: Gini coefficient by urban/rural localities in 2005/06.....	53
Table 6: Gini coefficient by regions in 2012/13	54
Table 7: Gini coefficient by ecological zones in 2012/13.....	55
Table 8: Gini coefficient by urban/rural localities in 2012/13.....	56
Table 9: Gini coefficient in 2005/06 and 2012/13 in Ghana	57
Table 10: Spatial differentials of income inequality for male headed households.....	58
Table 11: Spatial differentials of income inequality for female headed households.....	59
Table 12: Spatial differentials of income inequality for all heads of households.....	60
Table 13: Summary statistic of spatial differentials of income inequality in Ghana.....	61

CHAPTER ONE

INTRODUCTION

1.0 Background of study

Income inequality has become a hot button issue in the 21st century due to the popularity it has gained in the discussion of developmental issues. Income inequality is a broader concept than poverty. Unlike poverty, income inequality is defined over an entire population and does not only focus on the poor. It solely covers the degree of income disparities of a given population.

As poverty has become a subject of global interest, greater emphasis has been attached to the role of income distribution as well as growth in reducing poverty (Fosu, 2010; Ali & Thorbecke, 2000; Ravallion, 2001; Khan, 2009; Naschold, 2002). However, research has shown that regardless of the rate of economic growth it is less likely that economic growth can reduce poverty. Moreover, there is a growing consensus that extreme inequality in income can stunt growth itself (Birdsall, 2005).

Also spatial patterns of income inequality are often of importance in their own right as well as relevant to the political economy of policy-making (World Bank, 2006). According to Al-Hassan & Diao (2007), regional inequality deserves special attention as Ghana pursues the attainment of the Millennium Development Goals (MDGs) which aims at reducing poverty and hunger.

Following the increased attention income inequality has gained in recent times due to its impact on poverty and other economic issues, some studies have been conducted in Ghana to explore the relevance of the subject.

The results obtained from the Ghana Living Standards Survey conducted in 1991/1992 and 1998/1999 indicate a dramatic increase in inter-regional income differences. The Gini coefficient (a measure of income inequality) in 1991/1992 was 0.30 while that recorded in 1998/1999 was 0.60 depicting a 100% increase in income inequality (GSS, 2000). In northern Ghana income inequality increased by 25% compared to 9.7% in southern Ghana between the years 1991 and 2006. According to Shepherd, Gyimah-Boadi, Gariba, Plagerson & Musa (2005) the rising inequality between the southern and northern sectors of Ghana provides the most obvious example of inter-regional income inequality in Ghana. It was also discovered that while inequality exists between regions (inter-regional), they are more significant within regions (intra-regional) (Aryeetey, Owusu & Menash, 2009). At the district level, within district inequality was higher than inequality between districts. More importantly, the district level inequality shows a significant effect on poverty but with varying signs depending on the state of economic activity in the district (Annim, Mariwah & Sebu, 2012).

In spite of all these findings income inequality in the country is on the ascendency and poverty remains a threat to the developmental initiatives of the country. This study seeks to statistically analyze the spatial differentials of income inequality in Ghana with emphasis on the sex of the head of household and also presents the use of Bayesian estimation in the calculation of Gini coefficient which is a measure of income inequality.

1.1 Statement of the problem

Many researchers have devised and discussed various methods for calculating income inequality (Dasgupta, Sen & Starrett, 1973; Berribi & Sibling, 1985; Chotikapanich & Griffiths, 2000; Griffiths, 2008; Atkinson, 1970; Gastwirth, 1972; Mussard, Seyte & Terraza, 2003, Charles-Coll, 2011). However, the Gini coefficient has been recognized as the best single measure of income inequality (Morgan, 1962). Although the Gini coefficient can be computed using different methods, the most popular measure is derived from the Lorenz curve (De Maio, 2007). In order to obtain estimate of the Gini coefficient many empirical analysis employed the use of probability distributions which pertains to only income (Longford, Pittau, Zelli & Massari, 2010; Gastwirth, 1972; Prinsner, 2009; Aaberge, 1993). Meanwhile, the Lorenz curve is represented by a plot of the cumulative proportion of income (arranged in an ascending order) against the cumulative proportion of household size. In this study, the researcher used information on both income and household size in the estimation of the Lorenz curve and consequently the Gini coefficient through Bayesian parameter estimation.

Political analysts, political leaders and developmental organizations are faced with the obligation to come out with the next new idea after the Millennium Development Goals has expired in 2015. It was realized that the lack of focus on inequality was a key limitation of the MDGs and rightly this has become a major priority for the post 2015 agenda. Therefore, income inequality has become prominent in developmental issues which include the reduction of poverty.

Since 1980's, the poverty rate has been trending considerably downward in all regions of the world except in Sub-Saharan Africa where the first target of the Millennium Development Goal seems unlikely to be attained by 2015 (Fosu, 2010). Meanwhile, Hammer & Naschold (2000)

asserted that Sub-Saharan Africa and least developed countries will not be able to achieve the first Millennium Development Goal through growth alone.

Ghana, one of the strongest emerging economies in Sub-Saharan Africa has achieved a noticeable decline in poverty between the years 1992 and 2006. The proportion of the poor went down from 0.52 in 1992 to 0.39 in 2006. This decline in poverty resulted in about 1.7 million poor people moving out of poverty (GSS, 2007). The economic growth rates in Ghana between 1983 and 2000 led to an average growth rate of about 4.7%. After attaining an unprecedented high growth rate of about 15% in 2011 which was believed to be the fastest growth rate in the world at the time, the growth rate in the country has been dwindling consistently between 2011 and 2013. The growth rate in 2011 fell from 15% to 8.8% in 2012 and further declined to 7.1% in 2013 (GSS, 2014). Consequently, Ghana has not been able to sustain her high growth rate momentum attained in 2011. However, the expansion in the economy resulted in almost 50 percent reduction in poverty in 2006 from 1990 levels.

Therefore, Ghana stands out to be the first African country among other countries in the world that have the largest poverty reduction and most likely to meet the first MDG target of reducing poverty by half in 2015 in her continent (Easterling, Fox & Sands, 2008). Meanwhile, the fall in poverty has not been accompanied with a decline in income inequality (Coulombe & Wodon, 2007).

Ghana would have achieved the target for MDG 1 years earlier if the Gini index had reduced considerably. Thus, the persistent increase in income inequality makes poverty reduction difficult. And even growth is less effective in reducing poverty in high inequality countries (Mckay, 1997; Hammar & Naschold, 2000). This means that what matters for the complete eradication of

poverty does not rely solely on the rate of growth but also the distribution of income. Wodon (1999) confirmed that changes in income inequality will have large effect on poverty.

One important factor which influences the extent of income inequality is gender. Although, gender inequality is a global issue, it is made worse by income distribution. When women and men do not have equal access to the resources of a country or equal opportunities to take part in decision-making, there are direct and indirect economic and social consequences.

1.2 Objectives of the Study

The ultimate goals of the study are to develop an alternative measure of the Gini coefficient through the application of Bayesian estimation by using the Lorenz curve and to analyze the changes in income inequality in Ghana between 2005/06 and 2012/13 with emphasis on the sex of household head.

The specific objectives are as follows:

1. To establish the use of Bayesian parameter estimation in the calculation of Gini coefficients.
2. To construct confidence intervals for the population Gini coefficients.
3. To examine the differences in income inequalities among male headed households and also among female headed households between 2005/06 and 2012/13.
4. To determine the statistical significance of changes in income inequalities among heads of households in the various administrative regions, ecological zones and rural/urban localities in Ghana between 2005/06 and 2012/13.

1.3 Significance of the study

The main relevance of the study is to improve the targeting mechanism for policy implementation towards the alleviation of poverty in Ghana. Since income inequality affects not only poverty and financial decisions but all other sectors of the economy, this study will enable the government to make better economic decisions and will also inform government on the allocation of resources. This will help to resolve the issue of gender based inequalities in income in Ghana. Finally, this study seeks to provide a more statistical alternative way of computing Gini coefficient based on Bayesian method of estimation.

1.4 The organizational structure

The thesis is divided into five chapters; the first chapter provides an overview of the research topic and comprises: the background to the study; the statement of the problem; the objectives of the study and the justification of the study. Chapter two is concerned with the review of literatures in relation to the study. Chapter three provides an in-depth explanation of the methods of analysis. Chapter four focuses on the analysis of data and the discussion of results or findings. Chapter five provides summary of the study, conclusions with regards to the objectives of the study and some recommendations.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

The alleviation of poverty has become a priority in most countries since it serves as a benchmark for measuring the development of a country. With an increasing focus on addressing poverty issues, income inequality has attracted the attention of most researchers. This is because it has been realized that income inequality matters when it comes to making progress in poverty reduction (Addison & Cornia, 2001). Therefore, understanding the concept of income inequality and its measures are very paramount to this study.

2.1 Definition of inequality?

The word “inequality” triggers different ideas in the mind of a reader or a listener depending on his or her training. Inequality can be viewed as a departure from some idea of “equality” (Cowell, 2009). Inequality can be defined in terms of outcomes, opportunities or processes. For instance unequal outcomes in income may result in unequal opportunities in education or access to basic services like health.

It is also relevant to distinguish between vertical and horizontal inequalities. Vertical inequality refers to inequalities between individuals and horizontal inequality looks at inequalities between groups that may be based on age, gender, regions, ethnicity or religious affiliation. This research focuses on horizontal income inequality.

2.2 Measures of income inequality

The measures of inequality can be classified into two categories, namely the positive and normative measures. The positive measures of inequality describe the existing pattern of income distribution and summarize it into a single statistic while the normative measures explicitly base inequality on value judgment (Kanbur, 1984). Positive measures include range, relative mean deviation, Gini index etc. Some well-known normative measures are the Dalton measures and the Atkinson index. Sen (1973) also explained that many of the positive measures are special cases of the normative measures.

The simplest measure of income inequality sorts the population from the poorest to the richest and shows the percentage of income (or expenditure) attributable to each quintile or decile of the population. However, the Gini index (or coefficient) is by far the most popular measure for operationalizing income inequality. The reason may be attributed to the fact that it is easy to calculate and interpret. Hence, this study employs the use of Gini coefficient as a measure of income inequality.

2.3 The Gini coefficient

As discussed above, there are various methods for measuring income inequality. However, the most common measure used is the Gini coefficient. The Gini coefficient was developed in 1921 by an Italian sociologist and statistician, Corrado Gini. This coefficient measures the extent to which the income or expenditure among individuals or households within a country or a sub-population deviates from a perfectly equal distribution. The Gini coefficient assumes values between 0 (or 0%) and 1(or 100%) inclusive. While a Gini coefficient of 0 implies perfect

equality (which means equal share of income or expenditure), a Gini coefficient of 1 represents a state of perfect inequality (which means only one person or household has all the income). Hence, the closer a Gini coefficient is to zero the more equal the economy and the farther the Gini coefficient from zero the more unequal the economy.

2.4 Properties of the Gini coefficient

The Gini coefficient satisfies most of the desirable properties of a good measure of income (or consumption) inequality. Some of these properties satisfied by the Gini coefficient include:

- ❖ Mean independence: This means that if all income were doubled or change by the same proportion, the measure income inequality would not change.
- ❖ Population size independence: All other things being equal, if the population were to change, the measure of income inequality would not change.
- ❖ Symmetry: This implies that if there should be a swap in income between two people there should be no change in the measure of income inequality.
- ❖ Pigou-Dalton transfer sensitivity: Here, the transfer of income from rich to poor should reduce measured income inequality and a transfer from a poor person to a rich person should be recorded as a rise in inequality.
- ❖ Decomposability: Under this criterion, income inequality may be broken down into sub groups by any dimension such as between-group inequalities and within-group inequalities.
- ❖ Statistical testability: One should be able to test the significance of change in the index over time.

2.5 Measures of Gini coefficient

Different methods of estimation have been developed to calculate the Gini coefficient or index. A more mathematically convenient way of deriving the Gini coefficient developed by Dasgupta et al, (1973), is the one that estimates for a population homogeneous on the income values and that are indexed in a non-decreasing order, $y_{i-1} \leq y_i, i = 1, 2, \dots, n$ and the Gini coefficient is given by:

$$G = \frac{1}{n} \left(n+1 - 2 \left(\frac{\sum_{i=1}^n (n+1-i) y_i}{\sum_{i=1}^n y_i} \right) \right)$$

$$= \frac{2 \sum_{i=1}^n i y_i}{n \sum_{i=1}^n y_i} - \frac{n+1}{n} \quad (2.1)$$

Berrebi & Sibling (1985) showed that the Gini coefficient can be computed as:

$$G = \sum_{i=1}^n y_i \left[\frac{n-i}{n} - \frac{i-1}{n} \right] \quad (2.2)$$

where, n is the sample size and y_i is the proportion of total income earned by the individual or household whose income has the i^{th} rank in the income distribution, assuming that

$$y_1 \leq y_2 \leq \dots \leq y_i \leq \dots \leq y_n.$$

Chotikapanich & Griffiths (2000) suggested an alternative estimator of the Gini coefficient which is defined as:

$$G = \left(\sum_{i=1}^n x_i y_{i+1} \right) - \left(\sum_{i=1}^n x_{i+1} y_i \right) \quad (2.3)$$

where, x_i and y_i are the cumulative proportion of the population and income respectively. So that $x_n = 1$ and $y_n = 1$. The income variable Y_i are arranged in an increasing order such that $y_{i-1} \leq y_i$, $i = 1, 2, \dots, n$.

Another estimator of the Gini coefficient known as the Gini coefficient of mean difference which clearly proves that the Gini coefficient is a measure of dispersion is given by:

$$G = \frac{\frac{1}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^n |y_i - y_j|}{2\mu}, \quad i \neq j \quad (2.4)$$

where,

μ = average income

y_i = levels of income

$\sum_{i=1}^{n-1} \sum_{j=i+1}^n |y_i - y_j|$ = the sum of the absolute difference of all pairs of incomes.

The numerator of the estimator represents the mean absolute difference of all pairs of incomes.

2.6 Theoretical framework

This study can be explained within the framework of Lorenz curve and Gini coefficient. These theories play a central role in measuring the extent of income inequality. The Lorenz curve plots the cumulative proportion (percentage) of household income (or expenditure) which is sorted from lowest to highest income or expenditure (on the y-axis) against the cumulative proportion (or percentage) of household size (on the x-axis).. The Lorenz curve can also be derived based on the distribution of income. Here, let y be income and $f(y)$ be the continuous probability function of y . Also let $G(y)$ be the probability that a unit (household or individual) selected at random will have income less than or equal to y .

Then

$$G(y) = \int_0^y f(x)dx \quad (2.5)$$

The average income earned by these units is given by:

$$Q(y) = \int_0^y xf(x)dx \quad (2.6)$$

Now, the proportion of income earned by units whose incomes are less than or equal to y will be:

$$\eta(y) = \frac{Q(y)}{\mu} \quad (2.7)$$

where, μ is the average income of all the units.

A plot of $G(y)$ on the x-axis and $\eta(y)$ on the y-axis gives the Lorenz curve.

Generally, the Lorenz curve must satisfy the following conditions (Kakwani & Ponder, 1973):

1. If $G = 0$ then $\eta = 0$
2. If $G = 1$ then $\eta = 1$
3. $\eta < G$
4. The slope of the curve increases monotonically

The first and second conditions rule out the possibility of a household or unit earning zero or negative income. And the third and fourth conditions imply that the Lorenz curve must lie below the hypothetical line of equality.

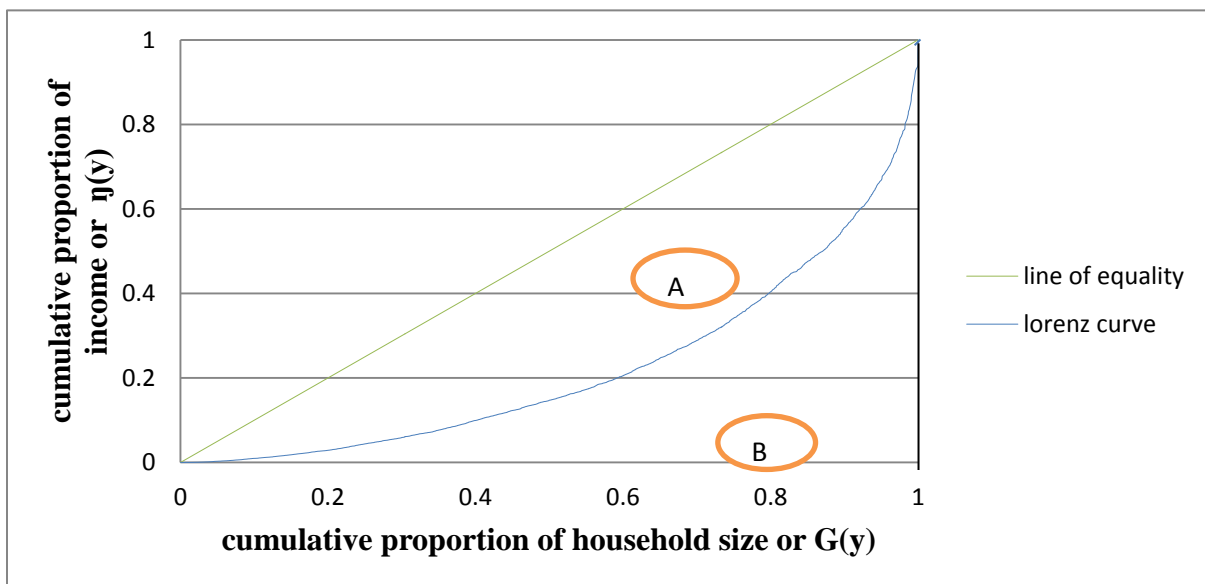


Figure 1: The Lorenz curve

The Gini coefficient is then computed as the area between the Lorenz curve and the line of equality (A), divided by the total area under the line of equality ($A + B$).

Thus the Gini coefficient is obtained by:

$$G = \frac{A}{A + B} = 2(0.5 - B) = 1 - 2B \quad (2.8)$$

where, $B = \int_0^1 L(x)dx$ and $L(x)$ is a function that represents the Lorenz curve.

There are alternative ways of computing the Gini coefficient. However, when the Lorenz curve is employed, the most commonly used method is the trapezium rule. Now, let (x_i, y_i) be a known point on the Lorenz curve then the area under the Lorenz curve (B) can be estimated using trapezoids and the Gini coefficient will be given as:

$$G = 1 - \sum_{i=1}^n (x_i - x_{i-1})(y_i + y_{i-1}) \quad (2.9)$$

Although the trapezium rule is simple, it is positive bias and consequently leads to a negative bias Gini coefficient (Fellman, 2012). Meanwhile, many researchers have proposed different parametric models for estimating the Lorenz curve (see Kakwani & Pongder, 1973; Thurow, 1970; Gaswirth, 1972; Bartels & van Metelen 1975; Kakwani, 1980). In most approaches a particular assumption about the statistical distribution of income is made and the parameters are estimated. Abdalla & Hassan (2004) compared the maximum likelihood and the non-linear least squares estimation technique and from their findings it was noticed that the non-linear least squares provided better and more reliable fit than the maximum likelihood method.

In this study, the researcher proposed the Bayesian estimation technique for estimating the Lorenz curve after fitting a regression model.

2.7 The decomposition of income inequality

Many empirical analysis of income inequality depends on measures of inequality which are decomposable (Bourguignon, 1979). The most common decomposable income inequality measures used are the generalized entropy measures (which include the Theil's T & Theil's L).

The generalized entropy measures are given by:

$$GE(\alpha) = \frac{1}{\alpha(\alpha-1)} \left[\frac{1}{n} \left(\sum_{i=1}^n \frac{y_i}{\bar{y}} \right)^\alpha - 1 \right], \quad \alpha \geq 0 \quad (2.10)$$

where,

α = the weight given to the distance between incomes at different levels.

n = the number of individuals or households in the sample.

$\left(\bar{y} = \frac{y}{n} \right)$ = the mean income per person or household.

Unlike the Gini coefficient the generalized entropy measure of income inequality assumes values from zero to infinity. When the inequality value is zero, this suggests an equal distribution of income and as the value increases, inequality in income distribution also increases.

The Theil's T and Theil's L which are also known as the mean log deviation measure are obtained when α is equal to 1 and 0 respectively. The Theil's L index is calculated as:

$$GE(0) = \left[\frac{1}{n} \sum_{i=1}^n \ln \left(\frac{\bar{y}}{y_i} \right) \right] \quad (2.11)$$

And the Theil's T is also given as:

$$GE(1) = \left[\frac{1}{n} \sum_{i=1}^n \frac{y_i}{\bar{y}} \ln \left(\frac{y_i}{\bar{y}} \right) \right] \quad (2.12)$$

These measures of income inequality allow the decomposition of total inequality into within-group inequality and between-group inequalities.

The Theil's T decomposition of income inequality into between-group and within-group inequalities is given as:

$$T = \sum_k \left(\frac{y_k}{y} \right) T_k + \sum_k \left(\frac{y_k}{y} \right) \ln \left[\left(\frac{y_k}{y} \right) \left(\frac{n}{n_k} \right) \right] \quad (2.13)$$

And that for Theil's L is given by:

$$L = \sum_k \left(\frac{n_k}{n} \right) L_k + \sum_k \frac{n_k}{n} \ln \left(\frac{\bar{y}}{\bar{y}_k} \right) \quad (2.14)$$

where,

n = represents the sample size of the entire population.

n_k = the sample size of the k th sub-group of the population.

y_k = the total income for the k th sub-group.

\bar{y} = the mean income of the entire population.

T_k = the value of $GE(1)$ for sub-group k .

L_k = the value of $GE(0)$ for the k^{th} sub-group.

For each of the decompositions above, the first term represents within-group inequality and the second term denotes between-group inequality.

This decomposition of income inequality helps in identifying the main source of income inequality in a given population. Usually, at least three-quarters of inequality in a country is due to within-group inequality and the remaining quarter to between-group differences.

According to Kim & Jeong (2003), Costa & Michelini has shown that the calibration of subpopulations to total inequality is more accurately done with the decomposition of the Gini index than the Theil's index. There has been different approach to the decomposition of the Gini index. Griffiths (2008) proved that the traditional decomposition of the Gini index and a new approach called the Dagum decomposition have identical component though their interpretation differs.

2.8 Empirical results

By using the Dagum Gini decomposition Ozkoc, Gurler & Ucdogruk (2011), found that the income inequality between gender, years, level of education and occupation in Turkey was significant meanwhile there was no significant difference in the levels of affluence between rural and urban areas. Lambert & Aronson (1993), throws more light by graphical analysis on the residual term that arises if the subgroup income ranges overlap in the decomposition of the Gini coefficient into within-groups and between-groups contributions. Senadza (2011) used the Gini

decomposition technique to investigate the effect of non-farm income on income inequality in rural Ghana. Meanwhile in Namibia, the Theil's T and Theil's L measures of entropy were used to ascertain the contribution of within-group inequality and between-group inequality to total income inequality. In this study within-group inequality was found to be the principal determinant of total inequality (Ekström, 1998). In Ghana, the generalized entropy measure of inequality was employed by Aryeetey et al (2009) and Annim et al (2012) to determine the within-region and between-region as well as within-district and between-district income inequalities. Most of the empirical researches reviewed were focused on the decomposition of income inequality into within-group and between group income inequalities. This study analyses the spatial differentials of income inequality with emphasis on sex of head of household. It also includes the construction of confidence intervals for the actual population Gini coefficients.

2.9 Definition of poverty?

Although poverty is different from inequality, they are related. Inequality focuses on the variations in living standards across a whole population. On the other hand, poverty deals with only those whose standard of living falls below a particular threshold which is referred as the poverty line. This threshold may be fixed in absolute terms (based on an externally determined norm, such as calorie requirements) or in relative terms (based on a fraction of the overall standard of living). Intuitively, relative poverty is more related to income inequality (Mckay, 2002).

2.10 The relationship between income inequality and poverty

The relation between income inequality and poverty has won the interest of many researchers in recent times (Addison & Cornia, 2001 and Abbas & Assane, 2006). This is because of the influence of income inequality in poverty reduction and also income inequality is essential to evaluate the benefit of a development strategy (Joshi & Gebremedhin, 2012). The most recent data showed that global poverty has been falling even as the world has become more unequal. This pattern was reflected in all the regions of the world except in Latin America and the Caribbean (LAC) where the reduction in poverty has been accompanied by a decline in income inequality. In Brazil, half of the total change in poverty between 2001 and 2009 was accredited to the reduction in income inequality. This result is even stronger for the case of extreme poverty (Lopez-Calva, 2012). After conducting a study to find out the long run and the short run impact of poor governance and income inequality through time series analysis for the year 1984-2008 in Pakistan, the results obtained confirmed positive relationship between income inequality and poverty both in the short run and long run. However, it was noticed that poor governance could have a significant impact on poverty in the long run but not in the short run (Akram, Wajid, Mahmood & Sarwar, 2011). In a research undertaken in India, the total change in poverty was decomposed into changes in the rise of mean income level and changes in the distribution of income. The results obtained indicated that, an increase in the mean income caused a decline in poverty but the effect was not fully realized and this was due to the changes in the distribution of income (Dhongde, 2002). This finding was also confirmed by Cheema & Sial (2012). They showed that the net growth elasticity of poverty is smaller than that of gross elasticity of poverty which suggests that some of the growth effect on poverty has been undermined by the rise in inequality. According to Gakuru & Mathenge (2012), in order to effectively address the poverty

situation in Kenya, the focus of government, economists and policy makers should not only be on growth but more importantly on the reduction of inequality. Poverty in Ethiopia has been a major issue of concern however there has not been a noticeable decrease in the levels of poverty. In an attempt to investigate the relationship between poverty, inequality and growth in rural Ethiopia, it was discovered that the inability of growth to reduce poverty substantially in the past was because of the increased inequality in the rural areas (Gelaw, 2009). By using a time series data of Middle East and North Africa (MENA), Ncube, Anyanwu & Hausken, (2013), realized that higher levels of inequality reduced the potential of growth to reduce poverty. Thus income inequality hinders the progress in attaining the prime goal of alleviating poverty and sustaining growth in the MENA region. High inequality can have undesirable political and social consequences (Alesina & Perotti, 1993).

2.11 Factors that influence income inequality

Most researchers have unveiled the negative impact of the rise in income inequality on poverty. In order to curb the influence of income inequality on poverty it is necessary to evaluate the causes of income inequality.

Charles-Coll (2011), classified the causes of income inequality into two categories, namely endogenous and exogenous. The endogenous or individual-specific causes of income inequality has to do with the characteristics or circumstances pertaining to an individual that can determine the future income as a result of influencing their comparative advantages either in the form of higher productivity or by the possession of scarce attributes which makes them comparatively more market- valuable and in a broader sense more socially competitive. The most basic group of

the endogenous causes of income inequality is the innate abilities of individuals. These may include intelligence, personality, charisma and physical attributes such as strength or skills, height and so on. A second group of the endogenous causes can be considered as necessary complement to the innate abilities. This involves the preferences or choices made by individuals. Most often the choices made by individuals irrespective of their inbuilt abilities can determine their level of income. Physical difference between individuals such as gender and race were also identified as endogenous causes of income inequality.

As the name suggests the exogenous causes are attributable to external factors that determine income levels. Some of the exogenous causes identified by Charles-Coll include land distribution, education, erroneous educational policies, labour market (or wage distribution), economic cycles, global recessions, globalization and intergenerational inequality. The land distribution is more associated to the rural setting where the possession of more lands means an increase in productivity. Education has been one of the major determinants of income level. In a society with poor access to education, the few inhabitants who obtain education may be allocated with positions that come with very high salaries. Most often, higher levels of education translate into higher salaries. This brings disparities in the income levels of individuals, thus influencing income inequalities. Erroneous educational policies could affect income inequality especially when demand characteristics of the labour market do not correspond to the skills acquired through education. Globalization through trade and financial liberalization has been found to be associated with income inequality. The global recession experienced at the beginning of 2008 rendered most people jobless bringing about differences in income levels. In some cases, income inequality occurs because of the influence of the decisions made by parents with regards to the choices they made in the past.

There are other possible sources of income inequality that have been investigated by researchers. In Germany, for instance between 2000 and 2006 it was realized that the unexpected increase in income inequality was due to changes in employment outcomes and in market returns as well as changes in the tax systems (Biewen & Juhasz, 2010). Unemployment does not only increase the risk of poverty but also contributes to inequality (Saunders, 2002). An increase in the student population in Denmark has led most of the youth to settle for low-wage jobs. This and other factors such as women engaging in full time jobs and family formulation caused the ascendancy of income inequality for 30 years (Neamtu & Westergaard-Nielsen, 2013). Bulř (2001), showed that the level of development, state employment, fiscal redistribution and price stability also serve as sources of income inequality. Reduction in inflation contributes immensely in lowering income inequality. This is because inflation appears to be one of the strongest factors fuelling income inequality especially in Africa (Anyanwu, 2011). Among other explanatory factors cyclical and structural changes in the labour market, the increasing relevance of capital income as well as the decreasing effectiveness of the public mechanism of income distribution improve income inequality (Schmid & Stein, 2013). In a country like Vietnam a substantial amount of total inequality was accounted for by spatial inequality, returns on education and the prevalence of white collar jobs (Heltberg, 2012). Corruption is another factor that worsens the levels of income inequality and also poverty in a country (Gupta, Davoodi & Alonso-Terme, 1998). The provision of education, health, transportation and financial services and also the execution of relevant development programs that will boost the income level of the population have been recognized as relevant for the fall in income inequality and the alleviation of poverty (Osahon & Osarobo, 2011). In Ghana, both internal and external remittances reduce the levels of

the depth and severity of poverty. However, there is a negative relationship between both types of remittances and income inequality (Adams, 2008).

2.12 Significance of spatial analysis

Spatial inequality is the uneven distribution of income and other socio-economic variables across different locations and regions (Aryeetey et al 2009).

Analyzing poverty and income inequality in a spatial context is becoming increasingly essential regarding the regional variations in their relationship (Joshi & Gebremedhin, 2012). Addressing ethnic and spatial inequalities is critical to poverty reduction for a number of reasons:

- ❖ First, between groups (or horizontal) inequalities make a large component of overall inequality within any country. A focus on only vertical inequality may obscure important differences among groups.
- ❖ Secondly, regional inequality in large industrializing countries, as well as in most developing and transition economies appears on the rise. If ethnic groups are geographically clustered industrialization or development may bypass groups that are not located in economic dynamic zones, intensifying poverty in the neglected areas.
- ❖ Third, inequalities between ethnic groups can lead to conflict, which is likely to affect development. Indeed, most conflicts today tend to have an ethnic dimension and are difficult to resolve.

- ❖ Fourth, horizontal or between-group inequalities are significant because in some situations it may not be possible to improve the position of individuals without tackling the position of groups.

CHAPTER THREE

METHODOLOGY

3.0 Introduction

This Chapter provides information on the source of data and presents a detailed discussion on the methods applied in the analysis of data.

3.1 Source of data

This study utilizes data extracted from the 2005/2006 and 2012/2013 Ghana Living Standards Survey. The Ghana Living Standards Survey (GLSS) is a nation-wide survey conducted by the Ghana Statistical Service (GSS). It provides vital information in assessing and monitoring the living conditions of households in Ghana. Ghana has conducted 6 rounds of the Living Standards Survey since 1987. The second, third, fourth, fifth and sixth rounds were carried out in 1988/1989, 1991/1992, 1998/1999, 2005/2006 and 2012/2013 respectively. Nationally representative samples of 8,687 households in 580 enumeration areas (EAs) as well as 18,000 households in 12,000 EAs were covered in GLSS5 and GLSS6 respectively. In both surveys, a two-stage stratified random sampling technique was adopted. At the first stage of sampling the EAs were allocated into the ten regions or strata using probability proportional to population size (PPS). This was followed by a complete listing of households in the selected EAs. At the second stage, a fixed take of 15 households were selected from each EA. This resulted in the nationwide sample sizes of 8,687 and 18,000 in GLSS5 and GLSS6 respectively. A comprehensive questionnaire which entailed key indicators such as health, education, demographic, migration, tourism, housing, non-farm

enterprises, employment, household income, consumption and expenditure among others was used in a face-to-face interview during both surveys.

3.2 Variables

The variables from GLSS 5 and GLSS 6 used in the analysis were nominal/gross household income, household size and sex of household head. The nominal income of the household constitutes income from rent, remittances, farm employment, non-farm employment and other sources. Household size also refers to the number of members in a household.

3.3 Data organization

Three spatial areas were considered in the study. These include the ten administrative regions, the ecological zones (Coastal, Forest and Savannah) and urban/rural localities.

The households sampled in these geographical areas were further grouped according to gender of household headship. The total sum of incomes earned by these sub-populations as well as the total sum of household sizes were calculated and the share (or proportion) of income and household size for the various households were also determined. The computed proportions of income for the head of households within each sub-population were sorted from the lowest to the highest with their corresponding proportion of household sizes. Then, the cumulative proportions of income and household size were computed. For each of the sub-populations, the cumulative proportions of income were plotted against the corresponding cumulative household sizes to obtain the Lorenz curve. In each instance, a regression model was fitted to the Lorenz curve and the coefficient

parameters were estimated through the application of Bayesian methods. Finally, the Gini coefficients for the various sub-populations are worked out. A detailed explanation of the procedures used in the derivation of the Lorenz curve and the Gini coefficient are discussed below.

3.4 Derivation of the Lorenz curve

Given a sample size of n , let M_k be the income of the k^{th} household and N_k be the corresponding household size, where $k = 1, 2, \dots, n$. Then, the proportion of the different levels of income will be given as:

$$Z_i = \frac{M_k}{\sum_{k=1}^n M_k}, \quad i = 1, 2, \dots, n \quad (3.1)$$

and the proportion of the household sizes will be:

$$S_i = \frac{N_k}{\sum_{k=1}^n N_k}, \quad i = 1, 2, \dots, n \quad (3.2)$$

Clearly, $\sum_{i=1}^n Z_i = 1$ and $\sum_{i=1}^n S_i = 1$

Now, let $Z_{(i)}$ be the i^{th} order statistic of the proportion of income such that:

$Z_{(1)}$ = the smallest of Z_1, Z_2, \dots, Z_n

$Z_{(2)}$ = the second smallest of Z_1, Z_2, \dots, Z_n .

.

.

$Z_{(j)}$ = the j^{th} smallest of Z_1, Z_2, \dots, Z_n

.

.

.

$Z_{(n)}$ = the largest of Z_1, Z_2, \dots, Z_n

So that, $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(n)}$

Let $S_{(i)}$ be the corresponding value of $Z_{(i)}$ ($i = 1, 2, \dots, n$), for the i^{th} household and let Y_i and W_i

be defined as follows:

$$Y_i = \sum_{j=1}^i Z_j \quad \text{and} \quad W_i = \sum_{j=1}^i S_j$$

Then clearly, $Y_n = 1$ and $W_n = 1$.

The functional relation between W_i and Y_i derived by $Y_i = L(W_i)$ is known as the Lorenz curve.

In this study, the Bayesian estimation method is used to determine the function L from the data.

Let A be the area between the Lorenz curve and the line of equality and let B be the area below the Lorenz curve. Then the Gini coefficient (G) is given as:

$$G = \frac{A}{A+B} = 1 - 2B \quad (3.3)$$

where,

$$B = \int_0^1 L(W)dw \quad (3.4)$$

A regression model $Y = L(W)$ will be fitted to the Lorenz curve and the parameters of the regression coefficient will be estimated using the Bayesian method of estimation.

3.5 The regression model

Because of the curvilinear nature of the Lorenz curve the regression model will take the form:

$$Y_i = \beta_1 W_i + \beta_2 W_i^2 + \beta_3 W_i^3 + \dots + \beta_p W_i^p + \varepsilon_i \quad (3.5)$$

The model does not have an intercept because the regression line will have to go through the origin.

3.5.1 Assumptions of the regression model

- ❖ The W_i 's are fixed (or known).
- ❖ The error term, ε_i is random and follows the normal distribution with mean zero and variance σ^2 (a constant). Thus $\varepsilon_i \sim N(0, \sigma^2)$.
- ❖ The Y_i 's are also random through ε_i and independent.
- ❖ W and Y are both continuous variables.

3.6 Bayesian estimation

Bayesian estimation begins with an assumed probability distribution on the parameter space, ϕ where the parameter(s) is(are) considered to be random. The probability distribution of the parameter(s) is(are) then updated based on the sample taken. Suppose the parameter space consists of a single parameter, θ . Then the distribution of the parameter θ on ϕ is called the prior distribution which reflects the researcher's belief about θ . The conditional probability of the data given the prior is called the likelihood. The posterior distribution then becomes the conditional distribution of the parameter after the sample data has been observed which integrates the prior and the sample information. Suppose, X is a random variable whose distribution depends on the parameter θ . Then the probability distribution of X is given as $f(x|\theta)$, let the prior distribution be $\pi(\theta)$. The posterior distribution of θ is obtained as follows:

$$f(\theta|x) = \frac{f(\theta, x)}{f_x(x)} = \frac{f(x|\theta)\pi(\theta)}{f_x(x)}$$

Because $f_x(x)$ is independent of θ , we can express the posterior distribution as proportional (\propto) to (Likelihood \times prior distribution).

That is,

$$f(\theta|x) \propto f(x|\theta)\pi(\theta) \quad (3.6)$$

Suppose X is a random vector from the distribution $f(x|\theta)$, then the posterior distribution is given as:

$$\begin{aligned} f(\theta/x) &= \frac{f(x_1, x_2, \dots, x_n, \theta)\pi(\theta)}{f(x_1, x_2, \dots, x_n)} \\ &= \frac{L(x_1, x_2, \dots, x_n | \theta)\pi(\theta)}{f(x_1, x_2, \dots, x_n)} \\ &\propto L(x_1, x_2, \dots, x_n | \theta)\pi(\theta) \end{aligned} \quad (3.7)$$

where, $L(x_1, x_2, \dots, x_n | \theta)$ is the joint probability distribution function of X called the likelihood function.

In the Bayesian setting, all the information about θ from the observed data and from the prior knowledge is contained in the posterior distribution. Therefore, the posterior distribution provides a more reliable estimation of θ than the prior. Hence, the Bayesian estimate of a parameter is the posterior mean. This knowledge can also be utilized in determining the Bayesian estimates for more than one parameter.

Now, from equation (3.5):

Let, $W_i = X_{1i}$, $W_i^2 = X_{2i}$, $W_i^3 = X_{3i}$, . . . , $W_i^p = X_{pi}$

So that equation (3.5) becomes:

$$\begin{aligned} Y_i &= \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_p X_{pi} + \varepsilon_i \\ &= \underline{\beta}' \underline{X}_i + \varepsilon_i \end{aligned} \quad (3.8)$$

where, $\underline{\beta}' = (\beta_1, \beta_2, \beta_3, \dots, \beta_p)$ and $\underline{X}_i = (X_{1i}, X_{2i}, X_{3i}, \dots, X_{pi})'$

Given the assumptions of the regression model, it follows that the Y_i 's are independent with each having a normal distribution with mean $\underline{\beta}' \underline{X}_i$ and variance σ^2 . So that the conditional distribution of the Y_i 's given $\underline{\beta}$ is:

$$f(Y_i | \underline{\beta}) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(Y_i - \underline{\beta}' \underline{X}_i)^2}; \quad |Y_i| \geq 0 \quad (3.9)$$

The likelihood function is given as:

$$\begin{aligned} L(Y | \underline{\beta}) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(Y_i - \underline{\beta}' \underline{X}_i)^2} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \underline{\beta}' \underline{X}_i)^2} \end{aligned} \quad (3.10)$$

Now, assuming the prior distribution of the random vector $\underline{\beta}$ is the multivariate normal distribution with mean vector $\underline{\mu}$ and covariance matrix, Σ . Thus, $\underline{\beta} \square N_p(\underline{\mu}, \Sigma)$

Then

$$\pi(\underline{\beta}) = \frac{1}{|\underline{\Sigma}|^{\frac{1}{2}}} \left(\frac{1}{\sqrt{2\pi}} \right)^p e^{-\frac{1}{2}Q(\underline{\beta})} \quad (3.11)$$

where, $Q(\underline{\beta}) = (\underline{\beta} - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{\beta} - \underline{\mu})$. Also, $\underline{\Sigma}^{-1} = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{p1} & \dots & a_{pp} \end{pmatrix}$ and $\underline{\mu} = (\mu_1, \mu_2, \mu_3, \dots, \mu_p)$

So the posterior distribution will be given as:

$$\begin{aligned} f(\underline{\beta}|Y) &\propto L(Y|\underline{\beta}) \times \pi(\underline{\beta}) \\ &\propto (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \underline{\beta}' \underline{X}_i)^2} \times (2\pi)^{-\frac{p}{2}} |\underline{\Sigma}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\underline{\beta} - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{\beta} - \underline{\mu})} \\ &= K \exp - \frac{1}{2} \left[\frac{1}{\sigma^2} \sum_{i=1}^n (Y_i - \underline{\beta}' \underline{X}_i)^2 + (\underline{\beta} - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{\beta} - \underline{\mu}) \right] \end{aligned} \quad (3.12)$$

where, K represents all terms that do not involve $\underline{\beta}$.

Now let

$$\begin{aligned} V &= \frac{1}{\sigma^2} \sum_{i=1}^n (Y_i - \underline{\beta}' \underline{X}_i)^2 \\ &= \frac{1}{\sigma^2} \sum_{i=1}^n \left(Y_i - \sum_{j=1}^p \beta_j X_{ji} \right)^2 \end{aligned} \quad (3.13)$$

Let, $i = 1$ and $j = 1, 2, 3$ then :

$$\begin{aligned} \left(Y_1 - \sum_{j=1}^3 \beta_j X_{j1} \right)^2 &= (Y_1 - \beta_1 X_{11} - \beta_2 X_{21} - \beta_3 X_{31})^2 \\ &= Y_1^2 - 2Y_1\beta_1 X_{11} - 2Y_1\beta_2 X_{21} - 2Y_1\beta_3 X_{31} + \beta_1^2 X_{11}^2 + 2\beta_1\beta_2 X_{11}X_{21} \\ &\quad + 2\beta_1\beta_3 X_{11}X_{31} + \beta_2^2 X_{21}^2 + 2\beta_2\beta_3 X_{21}X_{31} + \beta_3^2 X_{31}^2 \end{aligned} \quad (3.14)$$

From the above equation, equation (3.13) can be summarized as follows:

$$\begin{aligned} V &= \frac{1}{\sigma^2} \sum_{i=1}^n Y_i^2 + \frac{1}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^p \beta_j^2 X_{ji}^2 - 2 \frac{1}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^p Y_i \beta_j X_{ji} \\ &\quad + 2 \frac{1}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{p-1} \sum_{s=j+1}^p \beta_j \beta_s X_{ji} X_{si} \end{aligned} \quad (3.15)$$

Also let:

$$T = (\underline{\beta} - \underline{\mu})' \Sigma^{-1} (\underline{\beta} - \underline{\mu}) \quad (3.16)$$

Now, let

$$\underline{\beta} = (\beta_1, \beta_2, \beta_3), \quad \underline{\mu} = (\mu_1, \mu_2, \mu_3) \quad \text{and} \quad \Sigma^{-1} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Then equation (3.16) can be written as:

$$\begin{aligned}
& (\beta_1 - \mu_1 \quad \beta_2 - \mu_2 \quad \beta_3 - \mu_3) \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \beta_1 - \mu_1 \\ \beta_2 - \mu_2 \\ \beta_3 - \mu_3 \end{pmatrix} \\
&= a_{11}(\beta_1 - \mu_1)^2 + a_{21}(\beta_2 - \mu_2)(\beta_1 - \mu_1) + a_{31}(\beta_3 - \mu_3)(\beta_1 - \mu_1) \\
&\quad + a_{12}(\beta_1 - \mu_1)(\beta_2 - \mu_2) + a_{22}(\beta_2 - \mu_2)^2 + a_{32}(\beta_3 - \mu_3)(\beta_2 - \mu_2) \\
&\quad + a_{13}(\beta_1 - \mu_1)(\beta_3 - \mu_3) + a_{23}(\beta_2 - \mu_2)(\beta_3 - \mu_3) + a_{33}(\beta_3 - \mu_3)^2 \tag{3.17}
\end{aligned}$$

From equation (3.17), T (equation 3.16) can be generalized as:

$$\begin{aligned}
T &= \sum_{j=1}^p a_{jj}(\beta_j - \mu_j)^2 + 2 \sum_{j=1}^{p-1} \sum_{s \neq j}^p a_{js}(\beta_j - \mu_j)(\beta_s - \mu_s) \\
&= \sum_{j=1}^p a_{jj}(\beta_j^2 + \mu_j^2 - 2\beta_j\mu_j) + 2 \sum_{j=1}^{p-1} \sum_{s \neq j}^p (a_{js}\beta_j\beta_s - a_{js}\beta_j\mu_s - a_{js}\mu_j\beta_s + a_{js}\mu_j\mu_s) \\
&= \sum_{j=1}^p a_{jj}\beta_j^2 + \sum_{j=1}^p a_{jj}\mu_j^2 - 2 \sum_{j=1}^p a_{jj}\beta_j\mu_j + 2 \sum_{j=1}^{p-1} \sum_{s \neq j}^p a_{js}\beta_j\beta_s \\
&\quad - 2 \sum_{j=1}^{p-1} \sum_{s \neq j}^p a_{js}\beta_j\mu_s - 2 \sum_{j=1}^{p-1} \sum_{s \neq j}^p a_{js}\mu_j\beta_s + 2 \sum_{j=1}^{p-1} \sum_{s \neq j}^p a_{js}\mu_j\mu_s \tag{3.18}
\end{aligned}$$

Now

$$\begin{aligned}
 V + T = & \sum_{j=1}^p \left[\frac{1}{\sigma^2} \sum_{i=1}^n X_{ji}^2 + a_{jj} \right] \beta_j^2 + 2 \sum_{j=1}^{p-1} \sum_{s=j+1}^p \left[\frac{1}{\sigma^2} \sum_{i=1}^n X_{ji} X_{si} + a_{js} \right] \beta_j \beta_s \\
 & - 2 \sum_{j=1}^p \left[\frac{1}{\sigma^2} \sum_{i=1}^n X_{ji} Y_i + m \right] \beta_j + R
 \end{aligned} \tag{3.19}$$

where, $m = \sum_{j=1}^p a_{jk} \mu_k$, $j = 1, 2, 3, \dots, p$ and R is a constant term independent of β_j .

Clearly,

$$\pi(\underline{\beta} | Y) = K e^{-\frac{1}{2} Q_{\beta}(\beta)} \tag{3.20}$$

Hence the posterior distribution follows the multivariate normal distribution.

Let the matrix defining $Q_{\beta}(\beta)$ be \sum_{β}^{-1} , where \sum_{β} is a $p \times p$ matrix with an inverse \sum_{β}^{-1} .

The elements of \sum_{β}^{-1} are given as:

$$n_{jj} = \frac{1}{\sigma^2} \sum_{i=1}^n X_{ji}^2 + a_{jj}, \quad j = 1, 2, \dots, p \tag{3.21}$$

$$n_{js} = \frac{1}{\sigma^2} \sum_{i=1}^n X_{ji} X_{si} + a_{js}, \quad j \neq s \tag{3.22}$$

The constant term is a column vector $\underline{c} = (c_1, c_2, \dots, c_p)'$ of order p with the j^{th} element given as:

$$c_j = -2 \left[\frac{1}{\sigma^2} \sum_{i=1}^n Y_i X_{ji} + \sum_{k=1}^p a_{jk} \mu_k \right], \quad j = 1, 2, \dots, p \tag{3.23}$$

The mean vector of the posterior distribution is obtained as:

$$\underline{\mu}_{\beta} = -\frac{1}{2} \underline{\Sigma}_{\beta} \underline{C} \quad (3.24)$$

Meanwhile the Bayesian estimates of the regression parameters in equation (3.7) are given as:

$$\hat{\underline{\beta}}_{\sim} = \underline{\mu}_{\beta} \quad (3.25)$$

3.7 Estimating $\underline{\mu}$ and $\underline{\Sigma}$ of the prior distribution

The jackknife estimation approach was used to determine the estimates of $\underline{\mu}$ and $\underline{\Sigma}$ of the prior distribution. The jackknife or “take out one” method of estimation is an iterative process. Here, the parameters of interest are estimated using the entire sample then each data point is in turn eliminated from the sample to obtain the partial estimates of the parameters.

Let $\hat{\underline{\beta}}_i = (\hat{\beta}_{1i}, \hat{\beta}_{2i}, \dots, \hat{\beta}_{pi})'$; $i = 1, 2, \dots, n$ be the i^{th} jackknife estimate of the regression parameters from a given data set which consist of a response variable Y and predictor variables, X_1, X_2, \dots, X_p . Then the estimate of the mean vector $\underline{\mu}$ of the random vector $\underline{\beta}$ is given as:

$$\underline{\mu} = (\mu_1, \mu_2, \mu_3, \dots, \mu_p)'$$

$$\text{where } \mu_j = \frac{1}{n} \sum_{i=1}^n \hat{\beta}_{ji}; j = 1, 2, \dots, p; \quad (3.26)$$

and the estimate of the covariance matrix is given as.

$$a_{js} = \frac{1}{n-1} \sum_{i=1}^n (\hat{\beta}_{ji} - \hat{\mu}_j)(\hat{\beta}_{si} - \hat{\mu}_s); j=1,2,\dots,p \text{ and } s=1,2,\dots,p. \quad (3.27)$$

3.8 Model estimation

Supposing the estimated model is obtained as:

$$\hat{Y} = \hat{\beta}_1 W + \hat{\beta}_2 W^2 + \hat{\beta}_3 W^3 + \dots + \hat{\beta}_p W^p \quad (3.28)$$

Then equation (3.4) becomes:

$$\begin{aligned} \hat{B} &= \int_0^1 (\hat{\beta}_1 W + \hat{\beta}_2 W^2 + \hat{\beta}_3 W^3 + \dots + \hat{\beta}_p W^p) dw \\ &= \left[\frac{\hat{\beta}_1 W^2}{2} + \frac{\hat{\beta}_2 W^3}{3} + \frac{\hat{\beta}_3 W^4}{4} + \dots + \frac{\hat{\beta}_p W^{p+1}}{p+1} \right]_0^1 \\ &= \left[\frac{\hat{\beta}_1}{2} + \frac{\hat{\beta}_2}{3} + \frac{\hat{\beta}_3}{4} + \dots + \frac{\hat{\beta}_p}{p+1} \right] \end{aligned} \quad (3.29)$$

So the estimated Gini coefficient from equation (3.3) will be given as:

$$\hat{G} = 1 - 2 \left[\frac{\hat{\beta}_1}{2} + \frac{\hat{\beta}_2}{3} + \frac{\hat{\beta}_3}{4} + \dots + \frac{\hat{\beta}_p}{p+1} \right] \quad (3.30)$$

3.9 Model evaluation

In order to ascertain the goodness of the regression model [equation (3.28)] used in estimating the Lorenz curve based on the Bayesian estimation criterion, the R-square (coefficient of determination) and the Adjusted R-square will be computed. R-square is a statistical measure that determines the closeness of the data points to the fitted model while adjusted R-square penalizes for adding additional explanatory variable. Unlike R-square the adjusted R-square will only increase if an additional explanatory variable actually influences the dependent variable. Most often the R-square is greater than or equal to the adjusted R-square. With both measures, the closer they are to 1 the better the model and in most instances the closer they are to 0 the worse the model. Hence, R-square and Adjusted R-square falls between 0 and 1 inclusive.

The R-square is computed as:

$$R^2 = 1 - \frac{SSE}{SST}$$

$$= 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y}_i)^2} \quad (3.31)$$

where, SSE = Error Sum Squares

SST = Total Sum Squares

The adjusted R-square is also given as:

$$R_{adj}^2 = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - k - 1} \right] \quad (3.32)$$

where, n is the sample size and k is the number of explanatory variables.

3.10 Bootstrapping the Gini coefficient

The bootstrap method proposed by Bradley Efron is also known as resampling with replacement. Here the data points from the original data set are randomly selected with replacement to form new data sets of the same size as the original. Since each drawing is made from the entire data set, a simulated data set is likely to omit some points and have duplicates or triplicates of others. Therefore, it is essential to generate a large number of bootstrap samples (1000 or more) in order to minimize the variability that may occur given any analysis.

An estimated Gini coefficient was calculated from each of the bootstrap samples so that the standard error of the bootstrap distribution of the estimated Gini coefficients can be calculated using the following formula:

$$SE_{\hat{G}} = \sqrt{\sum_{i=1}^N (\hat{G}_i - \bar{G})^2 / N - 1} \quad (3.33)$$

where, \hat{G}_i = the Gini coefficient from the i^{th} bootstrap sample.

$$\bar{G} = \left(\sum_{i=1}^N \hat{G}_i \right) N^{-1} = \text{the mean of the Gini coefficient from the bootstrap samples.}$$

N = number of bootstrap samples.

The confidence interval for the true population Gini coefficient was obtained as:

$$\hat{G} \pm Z_{\alpha/2} SE_{\hat{G}} \quad (3.34)$$

Here, the distribution of the sampled estimated Gini coefficients from the bootstrap samples was presumed to be standard normal because of the large sample size.

3.11 Significance test for differences in Gini coefficient

To test the significant difference in income inequality between 2005/06 and 2012/13 given the various geographical areas with emphasis on gender the following hypotheses were considered.

Hypotheses:

H_0 : There is no significant difference in income inequality between 2005/06 and 2012/13.

H_1 : There is a significant difference in income inequality between 2005/06 and 2012/13.

Meanwhile, the two-way hypothesis test stated above was also applied to examine:

1. The difference in income inequality between 2005/06 and 2012/13 by regions, ecological zones and (urban/rural) localities with emphasis on the sex of head of household.
2. The significance of the change in income inequality between male headed household and also between female headed households in the entire country between 2005/06 and 2012/13

The null hypothesis (H_0) was rejected at the α -level of significance if the confidence interval for the actual sub-population Gini coefficient in 2005/06 does not overlap with that in 2012/13. Otherwise we failed to reject H_0 and conclude that there is no significant change in income inequality.

CHAPTER 4

ANALYSIS OF DATA AND DISCUSSION OF RESULTS

4.0 Introduction

This chapter deals with the analysis of data and presents a discussion on the results obtained. The chapter also provides descriptive analysis of the data used as well as the data processing tools employed.

4.1 Data processing

The statistical software that assisted in the analysis of data include R-Studio, SPSS, and MS Excel. The R-studio was used to write programs that derived the Gini coefficients. It was also used to obtain bootstrap samples for constructing confidence intervals for the true population Gini coefficients. The other softwares were used in performing other statistical tasks including the graphical presentation and tabulation of results.

4.2 Descriptive analysis of data

This section focuses on the sample sizes used in the analysis of data for the various sub-populations considered in the study. It also provides some descriptive statistics on household sizes and incomes of household heads from GLSS 5 and GLSS 6.

4.2.1 Sample sizes

The sample sizes used for the analysis from GLSS 5 and GLSS 6 were 8,375 and 16,551 correspondingly. This was as a result of cleaning the data by excluding household heads with either negative or no income.

Among the ten administrative regions in Ghana, the Ashanti region had the largest share of the sample size in both surveys while the Upper East region recorded the lowest share also in both surveys. The samples from the forest zone as well as the rural localities also had the highest count of interviewees in the two surveys. The proportion of male headed households in the entire sample was almost the same in both surveys (GLSS 5=71.83%, GLSS 6= 71.80%) and was more than twice the proportion of the female household heads. This proves that most of the households in Ghana are headed by males (See Appendix 1).

4.2.2 Descriptive statistics of income and household size

The two main variables considered in the analysis of data were income of household head and household size. This sub-section provides a discussion on some salient summary statistics on income of household head and household size from GLSS 5 and GLSS 6.

4.2.2.1 Income

Table 1: Summary statistic of household income

Statistic	GLSS 5			GLSS 6		
	Sex of household head		All	Sex of household head		All
	Male	Female		Male	Female	
Mean	1369.33	913.69	1240.99	16594.51	10188.57	14788.18
Error margin of mean	58.34	67.78	46.25	924.00	829.17	704.78
Maximum	52711.61	27328.96	52711.6	2184471	984446.69	2184469.3
Minimum	1.48	4.36	1.48	4.00	2.00	2.00
Standard deviation	2308.15	1678.87	2159.27	51388.44	28893.49	46257.61
Standard error	29.76	34.57	23.59	471.39	422.94	359.56

Source: Author's computation from GLSS 5 and GLSS 6

(Note: All income in GH¢. Also, 95% confidence interval for population mean = Mean \pm Error margin of mean)

From Table 1, it was realized that the average income earned by all male headed households in Ghana was significantly higher than that of the female headed households in both 2005/06 and 2012/13. This could mean that the male headed households in Ghana engage in higher paid jobs than their female counterparts. Also, the average income among the male headed household and the female headed household increased significantly between 2005/06 and 2012/13.

The maximum income earned by the male headed households was greater than that of the female headed households in both 2005/06 and 2012/13. Although, the minimum income earned by the female headed households was higher than that recorded among the male headed households in 2005/06, the female headed households had a lower minimum income in 2012/13.

It can be deduced from Table 1 that the income recorded for the male headed households during both GLSS 5 and GLSS 6 were more variable than that of the female headed households. It was also realized that the standard error for the average income earned by the male headed households

was higher than that of the female headed households. This means that the precision associated with the estimated average income with respect to the true average income was higher among the female headed households than their male counterparts.

4.2.2.2 Household size

Table 2: Summary statistic of household size

Statistic	GLSS 5			GLSS 6		
	Sex of household head		All	Sex of household head		All
	Male	Female		Male	Female	
Mean	5	4	5	5	4	5
Error margin of mean	0.08	0.08	0.06	0.05	0.06	0.04
Maximum	29	14	29	29	20	29
Minimum	1	1	1	1	1	1
Standard deviation	3.00	2.03	2.83	2.94	2.07	2.79
Standard error	0.04	0.04	0.03	0.03	0.03	0.02

Source: Author's computation from GLSS 5 and GLSS 6

(Note: 95% confidence interval for population mean = Mean \pm Error margin of mean)

From Table 2, the average household size for all male headed households was bigger than that of the female headed households during GLSS 5 and GLSS 6.

The maximum household size recorded for the male headed households was higher than their female counterparts during both surveys. The minimum household size for male and female headed households remained the same for both surveys.

Also, the data analyzed from both GLSS 5 and GLSS 6 showed that the variability in household size among the male headed households in Ghana was higher than that of the female headed households. The standard errors associated with estimated average household size was minimal

for both male and female headed households during GLSS 5 and GLSS 6 indicating that the estimated average household sizes were good representatives of the true average household sizes.

4.3 Estimation of Gini coefficient

This section provides detailed results on how the Gini coefficient and confidence interval for the population Gini coefficient for female headed households in the Central region were obtained from GLSS 6 based on the methodology explained in the previous chapter.

4.3.1 The Prior distribution

From equations (3.26) and (3.27), the jackknife estimates of μ and Σ were obtained as:

$$\hat{\mu} = \begin{pmatrix} 0.2480467 \\ -1.008396 \\ 1.558266 \end{pmatrix}$$

$$\text{and } \hat{\Sigma} = \begin{pmatrix} 6.9234E-07 & -2.7724E-06 & 2.4421E-06 \\ -2.7724E-06 & 1.1283E-05 & -1.0024E-05 \\ 2.4421E-06 & -1.0024E-05 & 8.9524E-06 \end{pmatrix}$$

The inverse of the covariance matrix, $\hat{\Sigma}^{-1}$ was also obtained as:

$$\hat{\Sigma}^{-1} = \begin{pmatrix} 555859545 & 357005940 & 248118830 \\ 357005940 & 246204129 & 178295357 \\ 248118830 & 178295357 & 132070580 \end{pmatrix}$$

4.3.2 The Posterior distribution

From equations (3.21) and (3.22) the covariance matrix defining $Q_{\beta}(\beta)$ of the posterior distribution was given as:

$$\hat{\Sigma}_{\beta}^{-1} = \begin{pmatrix} 556316725 & 357340285 & 248381656 \\ 357340285 & 246466956 & 178511507 \\ 248381656 & 178511507 & 132253938 \end{pmatrix}$$

The inverse of the covariance matrix above was obtained as:

$$\hat{\Sigma}_{\beta} = \begin{pmatrix} 6.8139E-07 & -2.7267E-06 & 2.4007E-06 \\ -2.7267E-06 & 1.1092E-05 & -9.8513E-06 \\ 2.4007E-06 & -9.8513E-06 & 8.7958E-06 \end{pmatrix}$$

Now the result obtained for the constant term defining $Q_{\beta}(\beta)$ of the posterior distribution from equation (3.23) was:

$$\hat{C} = -2 \begin{pmatrix} 164696877 \\ 118269319 \\ 87686861 \end{pmatrix}$$

The mean vector of the posterior distribution from equation (3.24) was obtained as:

$$\hat{\mu}_{\beta} = \begin{pmatrix} 0.2480468 \\ -1.0083958 \\ 1.5582662 \end{pmatrix}$$

Since, $\hat{\beta}_{\sim} = \mu_{\beta}$ the Bayesian parameter estimates of the regression model in equation (3.25) were also obtained as:

$$\hat{\beta}_{\sim} = \begin{pmatrix} 0.2480468 \\ -1.0083958 \\ 1.5582662 \end{pmatrix}$$

Hence the estimated Bayesian regression model which was used to represent the Lorenz curve from equation (3.23) can be written as:

$$\hat{Y} = 0.2480W - 1.0084W^2 + 1.5583W^3$$

This procedure was used to determine the estimated regression model for the other sub-populations covered in the study.

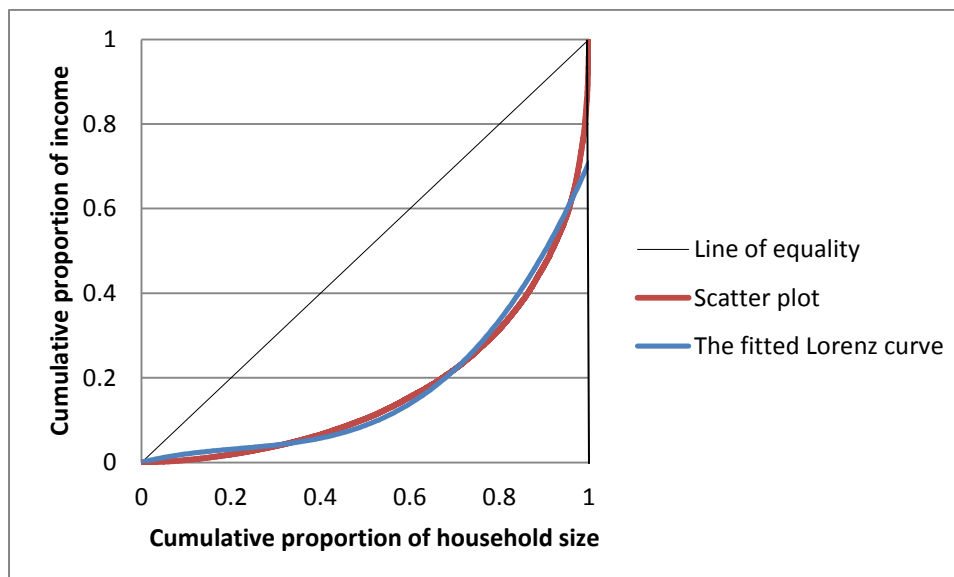


Figure 2: Graphical representation of the fitted Lorenz curve

The adjusted R-square for the fitted model was 0.99 which means that about 99% of the variation in the cumulative proportion of income has been explained by the model. This also means that the Bayesian estimation used provided good estimates of the regression parameters. The adjusted R-squares obtained for the other sub-population were all close to 1. Hence, with the various sub-populations the regression model fitted to the Lorenz curve was of the third order. (See Appendix 3).

After obtaining an estimate of the fitted model, the area under the Lorenz curve from equation (3.24) was obtained as:

$$\hat{B} = 0.1774$$

Finally, the estimated Gini coefficient for female headed households in the Central region was obtained as:

$$\hat{G} = 0.64512 \approx 0.645$$

The histogram and quantile plot for the Gini coefficients estimated from the 1000 bootstrap samples from Figure 3 provides a strong indication that the distribution of the sampled Gini coefficient from the bootstrap samples followed the Standard normal distribution.

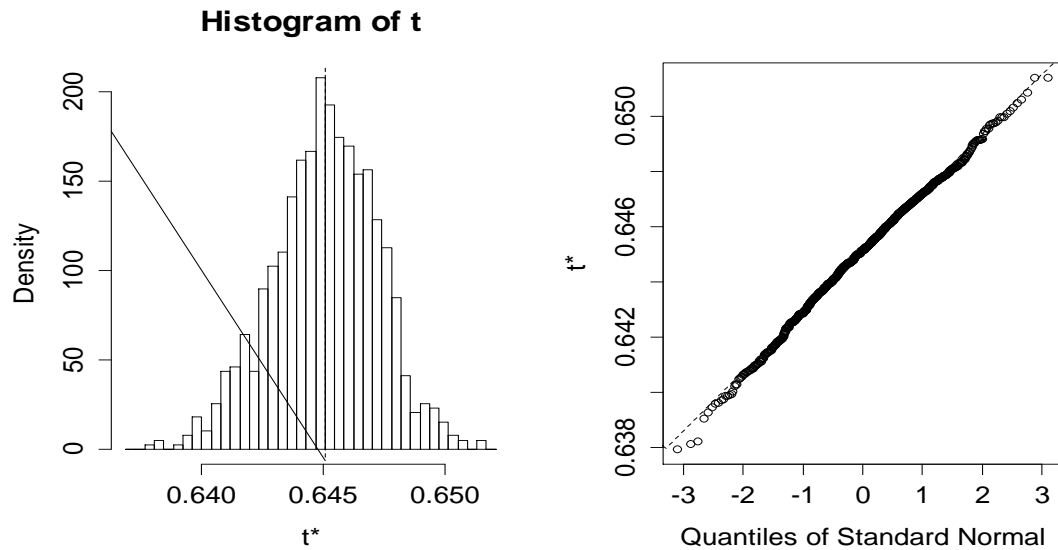


Figure 3: Graphical plot of Gini coefficient from bootstrap sample

Thus from the 1000 bootstrap samples the 95% confidence interval for the true population Gini coefficient was (0.641, 0.649). Hence, we are 95% confident that the true population Gini coefficient for female headed households in the Central region during GLSS 6 was between 0.641 and 0.649 given the sampled data.

The computed Standard error for the estimated Gini coefficients obtained from the bootstrap samples of the various sub-populations was low. This resulted in very narrow confidence intervals. This proved that the estimated Gini coefficients were good estimates of the population Gini coefficients.

4.4 Income inequality in Ghana during 2005/06

This section presents findings on the distribution of income in Ghana during 2005/06 by regions, ecological zones and urban/rural localities with emphasis on the sex of household headship.

4.4.1 Regions

Table 3: Gini coefficient by regions in 2005/06

Region	Male headed household		Female headed household		All heads of household	
	Gini	Error margin	Gini	Error margin	Gini	Error margin
Western	0.378	0.0024	0.487	0.0047	0.402	0.0020
Central	0.465	0.0041	0.555	0.0096	0.498	0.0037
Greater Accra	0.460	0.0031	0.518	0.0047	0.477	0.0026
Volta	0.427	0.0027	0.428	0.0037	0.427	0.0020
Eastern	0.448	0.0033	0.448	0.0029	0.450	0.0024
Ashanti	0.454	0.0022	0.487	0.0031	0.463	0.0018
Brong Ahafo	0.531	0.0043	0.465	0.0061	0.520	0.0031
Northern	0.529	0.0031	0.464	0.0102	0.523	0.0027
Upper East	0.546	0.0047	0.545	0.0153	0.546	0.0043
Upper West	0.492	0.0041	0.526	0.0153	0.491	0.0039

Source: Author's computation from GLSS 5

(Note: Confidence interval= Gini \pm Error margin)

The results obtained from Table 3 showed that among the male headed household from the administrative regions of Ghana, the Western region recorded the lowest share of income inequality. The region with the most unequal distribution of income was the Upper East region (Gini=0.546). This was also reflected among the households headed by men in the region. Among the female headed households within the various regions, it was found that most of the income generated in the Central region was shared by fewer households (Gini=0.555) compared to the rest. The extent of income inequality among households headed by women was the same in the

Western and Ashanti regions during GLSS 5. The region with the lowest income inequality across the administrative regions and among the female headed households (Gini=0.428) was the Volta region. Most of the regions in the Northern regions of Ghana had higher levels of inter-household income inequality and could explain why poverty rate in these regions is most glaring.

Also, the inequality in income recorded for female headed households in the Western, Central, Greater Accra, Ashanti and Upper West regions were significantly higher than that of their male counterparts. Although, the unequal distribution of income for the female headed households was higher than the male headed households in the Volta and Upper East regions, it was not statistically significant. The income inequalities for female headed households in the Brong Ahafo and Northern regions were significantly lower compared to that of the male headed households.

4.4.2 Ecological zones

Table 4: Gini coefficient by ecological zones in 2005/06

Ecological zone	Male headed household		Female headed household		All heads of household	
	Gini	Error margin	Gini	Error margin	Gini	Error margin
Coastal	0.469	0.0022	0.535	0.0035	0.491	0.0018
Forest	0.438	0.0014	0.470	0.0020	0.447	0.0010
Savannah	0.553	0.0018	0.507	0.0033	0.545	0.0016

Source: Author's computation from GLSS 5.

(Note: Confidence interval= Gini \pm Error margin)

It was revealed from Table 4 that the percentage of the unequal distribution of income among female headed households was more than 50% except in the forest zone. On the other hand inequality in income recorded among the male headed households in the various ecological zones was less than 50% except in the savannah zone. In all, the results obtained from the analysis

showed that the share of income in the savannah zone of Ghana had the highest share of income inequality.

The male headed households in the Coastal and Forest zones had income inequality significantly lower than the female headed households while the income inequality for the male headed households was significantly higher than their female counterparts in the Savannah zone.

4.4.3 Urban/ Rural localities

Table 5: Gini coefficient by localities in 2005/06

Locality	Male headed household		Female headed household		All heads of household	
	Gini	Error margin	Gini	Error margin	Gini	Error margin
Rural	0.511	0.0010	0.488	0.0018	0.503	0.0010
Urban	0.457	0.0018	0.504	0.0022	0.472	0.0014

Source: Author's computation from GLSS 5

(Note: Confidence interval= Gini \pm Error margin)

Table 5 showed that the disparity in the distribution of income among male household heads was more in the rural localities than in the urban localities. Meanwhile, the income gap among female headed households was bigger in the urban localities than in the rural localities. It was deduced from the information retrieved from the analysis that, the distribution of income was highly unequal among households in the rural localities than households in the urban localities.

In the rural localities, income inequality among the male headed households was significantly higher than that of the female headed households. In the urban communities, income inequality was rather significantly higher among the female headed households than the male headed households.

4.5 Income inequality in Ghana during 2012/13

This section also presents results from the analysis of data on income inequality in Ghana during 2012/13 for the various geographical areas with emphasis on the sex of heads of households.

4.5.1 Regions

Table 6: Gini coefficient by regions in 2012/13

Region	Male headed household		Female headed household		All heads of household	
	Gini	Error margin	Gini	Error margin	Gini	Error margin
Western	0.550	0.0026	0.584	0.0043	0.558	0.0022
Central	0.676	0.0049	0.645	0.0043	0.669	0.0033
Greater Accra	0.552	0.0029	0.619	0.0047	0.570	0.0024
Volta	0.613	0.0041	0.591	0.0051	0.611	0.0029
Eastern	0.521	0.0027	0.557	0.0047	0.532	0.0022
Ashanti	0.675	0.0041	0.642	0.0055	0.668	0.0031
Brong Ahafo	0.627	0.0041	0.664	0.0069	0.639	0.0033
Northern	0.620	0.0029	0.636	0.0065	0.620	0.0027
Upper East	0.701	0.0031	0.608	0.0080	0.689	0.0026
Upper West	0.689	0.0051	0.635	0.0090	0.686	0.0047

Source: Author's computation from GLSS 6
(Note: Confidence interval= Gini \pm Error margin)

The data analyzed from GLSS 6 showed that among the male headed households in the various regions of Ghana, those in the Upper East region had the highest proportion of income inequality. Interestingly, the Eastern region had the lowest uneven share of income among both male and female household heads. The variation in income among the female headed households in the Brong Ahafo region (Gini=0.664) was wider than those in the other regions. It was also discovered from Table 7 that the rates of income inequalities in the Eastern, Greater Accra and

Western region were below 60% but not lower than 50%. Just like the results obtained in GLSS 5, inequity in inter-household income inequality in Northern Ghana was quite higher than those recorded in the Southern part of the country.

It was also noted that income inequalities in the Western, Greater Accra, Eastern, Brong Ahafo and Northern regions were significantly higher among households headed by females than males. However, in the other regions income inequalities were significantly higher among male headed households than households headed by females.

4.5.2 Ecological zones

Table 7: Gini coefficient by ecological zones in 2012/13

Ecological zone	Male headed household		Female headed household		All heads of household	
	Gini	Error margin	Gini	Error margin	Gini	Error margin
Coastal	0.602	0.0024	0.629	0.0027	0.612	0.0018
Forest	0.619	0.0016	0.617	0.0024	0.620	0.0014
Savannah	0.664	0.0020	0.680	0.0039	0.666	0.0016

Source: Author's computation from GLSS 6
(Note: Confidence interval= Gini \pm Error margin)

From Table 7, it was noticed that very few people in the Savannah zone of Ghana have access to the income generated in the zone. This was because higher inequality in income was recorded among both male and female heads of households in the zone. The lowest proportions of income inequality among the female and male headed households across the ecological zones of Ghana were recorded in the forest and coastal zones respectively.

Inequality in income among male headed households in the Forest zone was higher than that among female headed households but not significant. Income inequalities among the female

headed households in the Coastal and Savannah zones were significantly higher than households headed by men.

4.5.3 Urban/ Rural localities

Table 8: Gini coefficient by localities in 2012/13

Locality	Male headed household		Female headed household		All heads of household	
	Gini	Error margin	Gini	Error margin	Gini	Error margin
Rural	0.601	0.0012	0.589	0.0022	0.598	0.0012
Urban	0.640	0.0020	0.645	0.0022	0.644	0.0016

Source: Author's computation from GLSS 6
(Note: Confidence interval= Gini \pm Error margin)

From Table 8, it was noticed that the male headed households in the urban localities recorded higher income inequalities than their female counterparts in the rural localities. On the other hand the distribution of income across households with women as heads was more unequal in the urban than in the rural.

It was also unveiled from the analyses of data that the unequal distribution of income among the male headed households in the rural localities was significantly higher than that of the female headed households. However, inequality in income in the urban localities was significantly higher than that of their male counterparts.

4.6 Summary statistic of income inequality for 2005/06 and 2012/13 in Ghana

Table 9: Gini coefficient in 2005/06 and 2012/13 in Ghana

Survey	Male headed household		Female headed household		All heads of household	
	Gini	Error margin	Gini	Error margin	Gini	Error margin
GLSS 5	0.509	0.0008	0.511	0.0014	0.507	0.0008
GLSS 6	0.649	0.0012	0.645	0.0016	0.647	0.0010

Source: Author's computation from GLSS 5 and GLSS 6
(Note: Confidence interval= Gini \pm Error margin)

Table 9 presents a summary of the income inequality observed in Ghana from GLSS5 and GLSS 6 based on all male and female headed households in Ghana. Consequently, it was noticed that the rate of income inequality among female headed households in the entire country was significantly higher than that of the male headed households in 2005/06. In 2012/13, the extent of income inequality among all female headed households in Ghana was significantly lower than their male counterparts.

4.7 Spatial differentials of income inequality

This section entails discussions on the spatial differentials of income inequality among heads of household in the geographical areas covered in the study between 2005/06 and 2012/13 with much emphasis on the sex of household heads.

Table 10: Spatial differentials of income inequality for male headed household

Geographical area	Male headed household				% change in Gini	
	GLSS 5 Gini	GLSS 5 95% CI	GLSS 6 Gini	GLSS 6 95% CI		
Region	Western	0.378	(0.376, 0.380)	0.550	(0.547, 0.553)	46%
	Central	0.465	(0.461, 0.469)	0.676	(0.671, 0.681)	45%
	Greater Accra	0.460	(0.457, 0.463)	0.552	(0.549, 0.555)	20%
	Volta	0.427	(0.424, 0.430)	0.613	(0.609, 0.617)	44%
	Eastern	0.448	(0.445, 0.451)	0.521	(0.518, 0.524)	16%
	Ashanti	0.454	(0.452, 0.456)	0.675	(0.671, 0.679)	49%
	Brong Ahafo	0.531	(0.527, 0.535)	0.627	(0.623, 0.631)	18%
	Northern	0.529	(0.526, 0.532)	0.620	(0.617, 0.623)	17%
	Upper East	0.546	(0.541, 0.551)	0.701	(0.698, 0.704)	28%
	Upper West	0.492	(0.488, 0.496)	0.689	(0.684, 0.694)	40%
Ecological zone	Coastal	0.469	(0.467, 0.471)	0.602	(0.600, 0.604)	28%
	Forest	0.438	(0.437, 0.439)	0.619	(0.617, 0.621)	41%
	Savannah	0.553	(0.551, 0.555)	0.664	(0.662, 0.666)	20%
Locality	Rural	0.511	(0.510, 0.512)	0.601	(0.600, 0.602)	18%
	Urban	0.457	(0.455, 0.459)	0.640	(0.638, 0.642)	40%

Source: Author's computation from GLSS 5 and GLSS 6

From Table 10, it was realized that between 2005/06 and 2012/13 income inequality among male household heads in the Ashanti region increased tremendously by almost 50% which was the highest across all the regions. The region with the lowest increase in income inequality among the male headed households was Eastern region. In all the regions, income inequality increased significantly among the male headed households in Ghana.

Income inequality also increased significantly among the male headed households in the various ecological zones. However, the percentage increase in income inequality among male headed households in the forest zone was almost twice of that recorded in the coastal and savannah zones.

Although the increase in income inequality was statistically significant among male headed households in the rural and urban localities, the percentage increase in income inequality among the male headed households was higher in the urban locality than in the rural locality.

Table 11: Spatial differentials of income inequality for female headed households

Geographical areas	Female headed household				% change in Gini	
	GLSS 5		GLSS 6			
	Gini	95% CI	Gini	95% CI		
Region	Western	0.487	(0.482, 0.492)	0.584	(0.580, 0.588)	20%
	Central	0.555	(0.545, 0.565)	0.645	(0.641, 0.649)	16%
	Greater Accra	0.518	(0.513, 0.523)	0.619	(0.614, 0.624)	19%
	Volta	0.428	(0.424, 0.432)	0.591	(0.586, 0.596)	38%
	Eastern	0.448	(0.445, 0.451)	0.557	(0.552, 0.562)	24%
	Ashanti	0.487	(0.484, 0.490)	0.642	(0.637, 0.647)	32%
	Brong Ahafo	0.465	(0.459, 0.471)	0.664	(0.657, 0.671)	43%
	Northern	0.464	(0.454, 0.474)	0.636	(0.630, 0.642)	37%
	Upper East	0.545	(0.530, 0.560)	0.608	(0.600, 0.616)	12%
	Upper West	0.526	(0.511, 0.541)	0.635	(0.626, 0.644)	21%
Ecological zone	Coastal	0.535	(0.531, 0.539)	0.629	(0.626, 0.632)	18%
	Forest	0.470	(0.468, 0.472)	0.617	(0.615, 0.619)	31%
	Savannah	0.507	(0.504, 0.510)	0.680	(0.676, 0.684)	34%
Locality	Rural	0.488	(0.486, 0.490)	0.589	(0.587, 0.591)	21%
	Urban	0.504	(0.502, 0.506)	0.645	(0.643, 0.647)	28%

Source: Author's computation from GLSS 5 and GLSS 6

The results of analysis recorded in Table 11 shows that households headed by women in the Brong Ahafo region had the largest rise in income inequality which was a little above 40%. The percentage increase in income inequality among female household heads in the Volta and Northern regions was close to 40%. The female headed households in the Upper East region had

the least percentage increase (12%) in income inequality among the ten regions. It can also be deduced from Table 11 that income inequality increased significantly among the female headed households between 2005/06 and 2012/13 in Ghana.

The female headed households with the smallest percentage increase in income inequality were in the coastal zone. Interestingly, the other two ecological zones had the same proportional increase in income inequality. The difference in income inequality among the female headed households in all the ecological zones was statistically significant.

Just as among the male headed household, income inequality was higher among female headed households in the urban locality than in the rural locality. However, the change in income inequality for the female headed households in both localities was significant.

Table 12: Spatial differentials of income inequality for all heads household

Geographical area	All heads of household				% change in Gini	
	Gini	GLSS 5 95% CI	Gini	GLSS 6 95% CI		
Region	Western	0.402	(0.400, 0.404)	0.558	(0.556, 0.560)	39%
	Central	0.498	(0.494, 0.502)	0.669	(0.666, 0.672)	34%
	Greater Accra	0.477	(0.474, 0.480)	0.570	(0.568, 0.572)	19%
	Volta	0.427	(0.425, 0.429)	0.611	(0.608, 0.614)	43%
	Eastern	0.450	(0.448, 0.452)	0.532	(0.530, 0.534)	18%
	Ashanti	0.463	(0.461, 0.465)	0.668	(0.665, 0.671)	44%
	Brong Ahafo	0.520	(0.517, 0.523)	0.639	(0.636, 0.642)	23%
	Northern	0.523	(0.520, 0.526)	0.620	(0.617, 0.623)	19%
	Upper East	0.546	(0.542, 0.550)	0.689	(0.686, 0.692)	26%
	Upper West	0.491	(0.487, 0.495)	0.686	(0.681, 0.691)	40%
Ecological zone	Coastal	0.491	(0.489, 0.493)	0.612	(0.610, 0.614)	25%
	Forest	0.447	(0.446, 0.448)	0.620	(0.619, 0.621)	39%
	Savannah	0.545	(0.543, 0.547)	0.666	(0.664, 0.668)	22%
Locality	Rural	0.503	(0.502, 0.504)	0.598	(0.597, 0.599)	19%
	Urban	0.472	(0.471, 0.473)	0.644	(0.642, 0.646)	36%

Source: Author's computation from GLSS 5 and GLSS 6

From Table 12, it was noticed that the regions with more than 40% increase in income inequality were the Ashanti and Volta regions. The Greater Accra and the Northern regions of Ghana had the same percentage increase (19%) in income inequality. The region with the smallest percentage increase (18%) in income inequality was the Eastern region. From Table 12, it was also realized that income inequality increased significantly in all the regions for all heads of household in Ghana between 2005/06 and 2012/13.

The proportional change in the inter-household income inequality between 2005/06 and 2012/13 was more pronounced in the forest zone of Ghana than in the coastal and savannah zones. The increase in income inequality in all the zones for all heads of household was significant.

From Table 10 and Table 11 it was noticed that the percentage increase in income inequality was higher among both male and female headed households in the urban locality than in the rural locality. This led to a higher proportional change in income inequality in the urban localities than in the rural localities for all heads of household. Irrespective of the difference in the percentage increase of income inequality in the various localities, income inequality increased significantly in both localities.

Table 13: Summary statistic of the spatial differentials of income inequality in Ghana

Sex of head of household	GLSS 5		GLSS 6		% change in Gini
	Gini	95% CI	Gini	95% CI	
Male	0.509	(0.508, 0.510)	0.649	(0.648, 0.650)	28%
Female	0.511	(0.510, 0.512)	0.645	(0.643, 0.647)	26%
All	0.507	(0.506, 0.508)	0.647	(0.646, 0.648)	28%

Source: Author's computation from GLSS 5 and GLSS 6

The results obtained from Table 13 shows that although income inequality increased more among male headed households than female headed households, the proportional change at each instance was statistically significant. It is therefore an indisputable fact that income inequality in Ghana is on the ascendency and remains a threat to national development.

CHAPTER 5

SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.0 Introduction

This chapter entails the summary of the objectives of the study and the methodology adopted in analyzing the data. It also presents conclusions drawn from the analysis of data and finally provides some recommendations.

5.1 Summary

The study focused on introducing a new method for computing Gini coefficient based on Bayesian estimation. It also analyzed the spatial differentials of household income inequality in Ghana between 2005/06 and 2012/13. The three spatial areas considered constitute the 10 administrative regions, the ecological zones and urban/rural localities. The household within these geographical areas were then grouped by gender of household headship.

Although, there are various methods for measuring the degree of income inequality, the Gini coefficient was adopted due to its popularity and simplicity. Meanwhile, the computation of the Gini coefficient also takes different forms. The study employs the use of Lorenz curve in determining the extent of income disparities among households in the various spatial areas considered.

The Bayesian estimation procedure was used to estimate the parameters of the regression model fitted to the Lorenz curve. A polynomial regression model of the third order without an intercept was found to best fit the Lorenz curve.

The bootstrap method was then employed to obtain the standard error of the original estimated Gini coefficients from the sample of Gini coefficients calculated from the bootstrap samples. These standard errors obtained for the different sub-populations were used in constructing a 95% confidence interval for their true population Gini coefficients.

The researcher was also interested in determining the statistical significance of the change in the Gini coefficients computed for the various group of households between 2005/06 and 2012/13. With this an overlap of the confidence intervals for the Gini coefficients in the different time periods imply that there has not been a significant change and the absence of an overlap indicated a significant change.

5.2 Conclusion

The findings from the study showed that the application of Bayesian estimation provided good estimates of the regression model fitted to the Lorenz curve. This was because the values obtained for R-square and adjusted R-square were above 90% for each sub-population considered.

The confidence intervals for the true population Gini coefficients derived for the various sub-populations were very narrow. This led to the conclusion that the estimated Gini coefficients for the different sub-population have very high precision.

A comparative analysis of the Gini coefficient revealed that the inter-household income inequality rose across all the spatial areas considered between 2005/06 and 2012/13. It was realized that the increase in unequal distribution of income among male headed households in the various administrative regions was considerably higher in the Western (46%), Central (45%), Volta

(44%), Ashanti (49%) and the Upper West (40%) regions of Ghana. The income gap among female headed household also worsened between 2005/06 and 2012/13 especially in the Brong Ahafo (43%), Volta (38%), Northern (37%) and the Ashanti (32%) regions of the country.

The income gap between the poor and the rich among the male headed households in the forest zone increased massively to about 41%. This kind tremendous increase was also noticed among female headed households in the savannah zone (34%).

The change in rural income inequality was observed to be positively higher among both male (40%) and female (28%) headed households between the two time periods considered in the study compared to their counterparts in the urban localities. Hence, income inequality in the urban localities has become a menace.

It was quite interesting to discover that the ascendancy in income inequality was more obvious among male headed households than female headed households in Ghana. It was discovered that not only was the spatial differentials of income inequality alarming but also statistically significant.

Meanwhile, there is a growing awareness that the continuous rise in income inequality stands in opposition to the development strategies of a country. The conclusions derived from the study point to an irrefutable fact that income inequality in Ghana is a “huge problem”. Government and policy makers are therefore charged with the responsibility of devising appropriate and feasible policies that will combat and eliminate the negative influence of income inequality.

5.3 Recommendation

The Bayesian estimation procedure employed in the study provided good estimates for the Lorenz curves with reference to the observed coefficients of determination. The confidence interval constructed for the true population Gini coefficient for the various sub-populations were narrow indicating that the estimated Gini coefficients are more precise. Unlike other parametric models that make use of the distribution of income in fitting the Lorenz curve, the regression model used in fitting the Lorenz curve relied on information on both household size and income. It is therefore recommended to researchers and policy makers to make use of the proposed method when measuring income inequality based on the Lorenz curve. It is also easier to construct confidence interval for the population Gini coefficient for grouped data when using the proposed method than other methods that depend on the distribution of income.

REFERENCES

- Aaberge, R. (1993). *Theoretical Foundation of Lorenz curve Orderings*. Discussion Paper No. 88, Research Department, Central Bureau of Statistics, Norway.
- Abbas, G., & Assane, D. (2006). The Poverty-Growth-Inequality Triangle Hypothesis: An Empirical Examination. *Journal of Policy Modelling*.
- Abdalla, M. I. & Hassan, Y. M. (2004). Maximum Likelihood Estimation of Lorenz Curves using Alternative Parametric Model. *Metodološki Zvezki*, Vol.1, No. 1, 109-118.
- Adams, Jr. H. R. (2008). *The Impact of Remittances on Poverty and Inequality in Ghana*. The World Bank Development Prospects Group, Development Economics Department and African Region, Policy Research Working Paper 4732.
- Addison, T., & Cornia, G. A. (2001). *Income Distribution Policies for Faster Poverty Reduction*. UNU-WIDER Discussion Paper No. 2001/9, World Institute for Development Economic Research, Helsinki.
- Akram, Z., Wajid, S., Mahmood, T., & Sarwar, S. (2011). Impact of Poor Governance and Income Inequality on Poverty in Pakistan. *Far East Journal of Psychology and Business*, Vol. 4 No. 3.
- Alesina, A., & Perotti, R. (1993). Income Distribution, Political Instability and Investment, *European Economic Review*, pp. 1203-1228.
- Al-Hassan, R. M. & Diao, X. (2007). Regional disparities in Ghana: Policy options and public investment implication, in *IFPRI Discussion Paper 00693*, Washington: IFPRI.
- Ali, A., & Thorbecke, E. (2000). The State and Path of Poverty in Sub-Saharan Africa: Some Preliminary Results. *Journal of African Economics* 9, 9-41.

- Annim, S., Mariwah, S., & Sebu, J. (2012). *Spatial Inequality and Household Poverty in Ghana*. BWPI Working Paper 161, Economic Systems, Vol. 36, No.4, 2012.
- Anyanwu, J. C. (2011). *International Remittance and Income Inequality in Africa*, Working Paper Series No. 135, African Development Bank, Tunis, Tunisia.
- Aryeetey, E., Owusu, G., & Mensah, J. E. (2009). *An Analysis of Poverty and Regional Inequalities in Ghana*. GDN Working Paper Series. Working paper No. 27.
- Atkinson, A. B. (1970). On the Measurement of Inequality. *Journal of Economic Theory*, pp. 244-263.
- Bartels, C. P. A. & van Metelen, H. (1975). Alternative Probability Density Functions of Income. Vrije University Amsterdam: *Research memorandum 29*, 30 pages.
- Berribi, M., & Sibling, J. (1985). Income Inequality Indices and Deprivation: A Generalization, *Journal of Economics* 99, 807-810.
- Biewen, M., & Juhasz, A. (2010). *Understanding Rising Income Inequality in Germany*. IZA Discussing Paper No. 5062.
- Birdsall, N. (2005). The World is not Flat: Inequality and Justice in our Global Economy: *WIDER Lecture 9*, United Nations University- WIDER, Helsinki.
- Bourguignon, F. (1979). Decomposable Income Inequality measures. *Econometrica*, Vol. 47, No.4 (Jul., 1979), pp.901-920.
- Buliř, A. (2001). *Income Inequality: Does Inflation Matter?* IMF Staff Papers, Vol. 48 No. 1.

- Charles-Coll, A. J. (2011). Understanding Income Inequality: Concept, Causes and Measurement. *International Journal of Economics and Management Sciences*. Vol.1, No. 3, 2011, pp. 17-28.
- Cheema, R. A., & Sial, H. M. (2012). Poverty, Income Inequality, and Growth in Pakistan: A Pooled Regression Analysis. *The Lahore Journal of Economics*, 17 : 2 (Winter 2012): pp. 137-157.
- Chotikapanich, D., & Griffiths, W. (2000). *Posterior Distribution for the Gini Coefficient Using Group data*.
- Coulombe, H., & Wodon, Q. (2007). *Poverty, Livelihoods and access to Basic Services in Ghana*. Partial and Preliminary draft for review.
- Cowell, A. F. (2009). *Measuring Inequality*. Oxford University Press.
- Dasgupta, P., Sen, A., & Star-rett, D. (1973). Notes on the Measurement of Inequality. *Journal of Economics Theory*, 6:180-7.
- De Maio, F. G. (2007). Income Inequality Measures. *Journal of Epidemiology and Community Health*, 61(10), 849-852. doi: 10.1136/jech2006.052969.
- Dhongde, S. (2002). *Measuring the Impact of Growth and Income Distribution on Poverty in India*. University of California.
- Easterling, T., Fox, W. J., & Sands, B. F. (2008). *Factors affecting Economic Growth in Ghana: Bases for a New USAID Approach to Economic Growth*. Contract EEM-I-02-07-00003-02, Sibley International.

- Ekström, E. (1998). *Income Distribution and Labour Market Discrimination: A case study of Namibia*. Research Institute of Industrial Economics (IUI). IUI Working Paper Series No. 502.
- Fellman, J. (2012). Estimation of Gini coefficients using Lorenz curves. *Journal of Statistical and Econometric Method*, Vol.1, No.2, 31-38.
- Fosu, K. A. (2010). *Income Distribution and Growth's Ability to Reduce Poverty: Evidence from Rural and Urban African Economies*. Working Paper No.2010/92.
- Gakuru, R., & Mathenge, N. (2012). *Poverty, Growth, and Income Distribution in Kenya: A SAM Perspective*. AGRODEP Working Paper 0001.
- Gastwirth, J. L. (1972). The estimation of the Lorenz curve and Gini index. *The Review of Economics Statistics*, Vol. 54, No. 3, pp. 306-316.
- Gelaw, F. (2009). *The Relationship between Poverty, Inequality, and Growth in the Rural Ethiopia: Micro Evidence*. Contribution Paper prepared for presentation at the International Association of Agricultural Economists Conference, Beijing, China.
- Ghana Statistical Service (2000). *Ghana Living Standards Survey: Report of the fourth round (GLSS 4)*, Accra, Ghana.
- Ghana Statistical Service (2007). *Pattern and Trends of Poverty in Ghana 1991-2006*.
- Ghana Statistical Service (2014). *Ghana Living Standards Survey Round 6: Poverty Profile in Ghana (2005-2013)*
- Griffiths, W. (2008). *On Dagum's Decomposition of the Gini Coefficient*. The University of Melbourne. Research Paper No. 1054.

- Gupta, S., Davoodi, H., & Alonso-Terme, R. (1998). *Does corruption affect income inequality and Poverty?* IMF Working Paper No.98/76 (Washington: International Monetary Fund).
- Hammer, L., & Naschold, F. (2000). Attaining the International Development Targets: Will Growth be Enough? *Development Policy Review* 18, 11-36.
- Heltberg, R. (2012). Causes of Inequality in Vietnam. *Institute of Economics*, University of Copenhagen.
- Joshi, S. & Gebremedhin T. (2012). A Spatial Analysis of Poverty and Income Inequality in the Appalachian Region. *The Journal of Rural and Community Development*, 7(2), 118-130.
- Kakwani, N. C. (1980). *Income Inequality and Poverty Methods of Estimation and Policy Applications*. Oxford University Press.
- Kakwani, N., & Pongder, N. (1973). On the Estimation of Lorenz curves from Grouped observations. *International Economic Reviews*, Vol. 14, No. 2, 278-292.
- Kanbur, S. M. R. (1984). The Measurement and Decomposition of Inequality and Poverty
- Khan, H. M. (2009). Governance, Growth and Poverty Reduction. DESA Working Paper No.75 ST/ESA/2009/DWP/75.
- Kim, E., & Jeong, H. Y. (2003). Decomposition of Regional Income Inequality in Korea. *The Review of Regional Studies*, Vol. 33, No. 3, 2003, pp. 313-327.
- Lambert, P. J., & Aronson, J. R. (1993). Inequality Decomposition Analysis and the GINI Coefficient revisited. *Economic Journal*, Vol. 103, pp. 1221-27.

- Longford, N. T., Pittau, M. G., Zelli, R., & Massari, R. (2010). *Measures of Poverty and Income Inequality in the Countries and Regions of EU*. ECINEQ, Working Paper 182, Society for the study of Economic Inequality.
- Lopez-Calva, F. L. (2012). *Declining Income Inequality in Brazil: The Proud Outlier. Inequality in Focus*, The World Bank.
- Mckay, A. (1997). Poverty Reduction through Economic Growth: Some Issues. *Journal of International Development* 9, 665-673.
- Mckay, A. (2002). *Defining and Measuring Inequality*. Overseas Development Institute and University of Nottingham, Briefing Paper No. 1(1 of 3).
- Morgan, J. (1962). The Anatomy of Income Distribution. *The Review of Economics and Statistics*, 44, pp. 270-283.
- Mussard, S., Seyte, F., & Terraza, M. (2003). Decomposition of Gini and the Generalized Entropy Inequality Measures. *Economic Bulletin*, Vol. 4, No. 7, pp. 1-6.
- Naschold, F. (2002). *Why Inequality Matters for Poverty*. Overseas Development Institute. Briefing Paper No. 2 (2 of 3).
- Ncube, M., Anyanwu, J., & Hausken, K. (2013). *Inequality, Economic Growth, and Poverty in the Middle East and North Africa (MENA)*, Working Paper Series No. 195 African Development Bank, Tunis, Tunisia.
- Neamtu, I., & Westergaard-Nielsen (2013). *Sources and Impact of Rising Inequality in Denmark*. Center for Corporate Performance, Aarhus University.
- Osahon, S., & Osarobo, K. A. (2011). Poverty and Income Inequality in Nigeria: An Empirical Assessment. *JORIND* 9(2), ISSN 1596- 8308.

- Ozkoc, H. H., Gurler, K. O., & Ucdogruk, S. (2011). Decomposition of Income Inequality in Turkey. *Current Research Journal of Social Sciences* 3(2): 96-103, 2011.
- Prisner, E. (2009). Income Distribution, Lorenz curve and Gini index: A Complex Calculus Project for the Classroom. Franklin College, Switzerland.
- Ravallion, M. (2001). Growth, Inequality and Poverty: Looking beyond Averages. *World Development* 29, 1803-1815.
- Saunders, P. (2002). *The Direct and Indirect Effects of Unemployment on Poverty and Inequality*. The Social Policy Research centre, Discussion Paper No. 118.
- Schmid, K. D., & Stein, U. (2013). Explaining the Rising Income Inequality in Germany, 1991-2010. *Macroeconomic Policy Institute*, pp.32.
- Sen, A. (1973). Inequality and Unemployment : Some Conceptual Issues in Measurement. *Economic and Political Weekly*, Vol. 8, No. 31/33, Special Number (Aug. 1973), pp.1457-1464.
- Senadza, B. (2011). Does Non-farm Income Improve or Worsen Income Inequality? Evidence from Rural Ghana. *African Review of Economics and Finance*, Vol. 2, No. 2, June 2011.
- Shepherd, A., Gyimah-Boadi, E., Gariba, S., Plagerson, S., & Musa, W. A. (2005). *Bridging the North-South divide in Ghana*, Background Paper for the UNDP 2006 World development report.
- Thurow, L. C. (1970). Analysing the American Income Distribution. *American Economic Review*, 60, 261-269.
- Wodon, Q. T. (1999). *Growth, Poverty and Inequality: A regional panel for Bangladesh*, Policy Research Working Paper No. 2072, World Bank, Washington, DC.

World Bank, (2006). World Development Report 2006. Equity and Development. Oxford University Press.

Appendix 1

Table 1A and 2B provides information on the sample sizes for the various sub-populations for GLSS 5 and GLSS 6.

Table 1A: Sample composition for heads household from GLSS 5

Geographical areas		Sex- Head of household					
		Male		Female		Total	
		n	%	n	%	n	%
Region	Western	578	71.71	228	28.29	806	100
	Central	424	63.28	246	36.72	670	100
	Greater Accra	836	70.02	358	29.98	1194	100
	Volta	488	70.01	209	29.99	697	100
	Eastern	582	66.14	298	33.86	880	100
	Ashanti	1002	66.45	506	33.55	1508	100
	Brong Ahafo	505	66.10	259	33.90	764	100
	Northern	677	88.15	91	11.85	768	100
	Upper East	493	84.13	93	15.87	586	100
	Upper West	431	85.86	71	14.14	502	100
	Total	6016	71.83	2359	28.17	8375	100
Ecological zone	Coastal	1655	66.73	825	33.27	2480	100
	Forest	2298	67.45	1109	32.55	3407	100
	Savannah	2063	82.92	425	17.08	2488	100
	Total	6016	71.83	2359	28.17	8375	100
Locality	Rural	3693	75.23	1216	24.77	4909	100
	Urban	2323	67.02	1143	32.98	3466	100
	Total	6016	71.83	2359	28.17	8375	100

Source: Author's computation from GLSS 5

Table 2A: Sample composition for heads household from GLSS 6

Geographical areas		Sex-Head of household					
		Male		Female		Total	
		n	%	n	%	n	%
Region	Western	1223	72.28	469	27.72	1692	100
	Central	962	61.75	596	38.25	1558	100
	Greater Accra	1310	70.05	560	29.95	1870	100
	Volta	1035	66.09	531	33.91	1566	100
	Eastern	1240	68.93	559	31.07	1799	100
	Ashanti	1287	65.70	672	34.30	1959	100
	Brong Ahafo	1113	70.35	469	29.65	1582	100
	Northern	1498	88.38	197	11.62	1695	100
	Upper East	1084	75.28	356	24.72	1440	100
	Upper West	1132	81.44	258	18.56	1390	100
	Total	11884	71.80	4667	28.20	16551	100
Ecological zone	Coastal	2535	64.82	1376	35.18	3911	100
	Forest	4711	68.21	2196	31.79	6907	100
	Savannah	4638	80.90	1095	19.10	5733	100
	Total	11884	71.80	4667	28.20	16551	100
Locality	Rural	7072	76.20	2209	23.80	9281	100
	Urban	4812	66.19	2458	33.81	7270	100
	Total	11884	71.80	4667	28.20	16551	100

Appendix 2

The R program used for the data analysis is as follows:

```
install.packages("bootstrap")
install.packages("boot")
library("bootstrap")
library("boot")
xdata<- read.csv("C:/Users/Benedicta PC/Desktop/REGIONS_6/ez1.csv")
xdata<- data.frame(xdata)
xdata

# MSE calculation
ginifxn<- function(xdata){
  regmodel<- lm(CUMINC~CUMSIZE+CUMSIZE2+CUMSIZE3-1, data=xdata)
  regmodel
  sm<- summary(regmodel)
  sm

  mse<- function(sm){
    mse<-mean(sm$residuals^2)
    return(mse)
  }
  MSE<- mse(regmodel)
  MSE
```

Jackknife estimation of μ and Σ of the prior distribution

```

DF<- xdata
model.lm<- formula(CUMINC~CUMSIZE+CUMSIZE2+CUMSIZE3-1)
theta<- function(x, xdata, coefficient){
  coef(lm(model.lm, data=xdata[x,]))[coefficient]}
results<- jackknife(1:nrow(xdata), theta, xdata=DF, coefficient="(Intercept)")

# The following function calculates all the coefficients
jackknife.apply <- function(x, xdata, coefs){
  sapply(coefs,function(coefficient) jackknife(x, theta, xdata=xdata,
      coefficient=coefficient),simplify=F)
}

# now jackknife.apply() can be called
results <- jackknife.apply(1:nrow(xdata), DF, c("CUMSIZE","CUMSIZE2","CUMSIZE3"))
results

V1<- results$CUMSIZE$jack.values
V2<- results$CUMSIZE2$jack.values
V3<- results$CUMSIZE3$jack.values

Final<- data.frame(V1,V2,V3)
Final

meanofCUMSIZE=mean(Final$V1)
meanofCUMSIZE2=mean(Final$V2)
meanofCUMSIZE3=mean(Final$V3)

```

```
meanofCUMSIZE
meanofCUMSIZE2
meanofCUMSIZE3

#the meanvector - mu_k
MeanVector<- c(meanofCUMSIZE,meanofCUMSIZE2,meanofCUMSIZE3)
MeanVector

d1<- V1-meanofCUMSIZE
d2<- V2-meanofCUMSIZE2
d3<- V3-meanofCUMSIZE3

Dmatrix<-data.frame(d1,d2,d3)
Dmatrix

m11<- sum(Dmatrix$d1*Dmatrix$d1)/nrow(xdata)
m22<- sum(Dmatrix$d2*Dmatrix$d2)/nrow(xdata)
m33<- sum(Dmatrix$d3*Dmatrix$d3)/nrow(xdata)
m12<- sum(Dmatrix$d1*Dmatrix$d2)/nrow(xdata)
m13<- sum(Dmatrix$d1*Dmatrix$d3)/nrow(xdata)
m23<- sum(Dmatrix$d2*Dmatrix$d3)/nrow(xdata)

xx<- c(m11,m12,m13)
xy<- c(m12,m22,m23)
yy<- c(m13,m23,m33)
```

```
gg<- data.matrix(data.frame(xx,xy,yy))
```

```
gg
```

Estimating $\hat{\Sigma}^{-1}$

```
#the a_sub i's
```

```
ggInverse<-solve(gg)
```

```
ggInverse
```

```
ggInverseMSE<-ggInverse
```

```
ggInverseMSE<-data.matrix(data.frame(ggInverseMSE))
```

```
ggInverseMSE
```

Estimating the elements of \sum_{β}^{-1} of the posterior distribution

```
head(xdata)
```

```
sumofsquaresCUMSIZE<- sum(xdata$CUMSIZE^2)/MSE
```

```
sumofsquaresCUMSIZE2<- sum(xdata$CUMSIZE2^2)/MSE
```

```
sumofsquaresCUMSIZE3<- sum(xdata$CUMSIZE3^2)/MSE
```

```
sumofCUMSIZE12<-sum(xdata$CUMSIZE*xdata$CUMSIZE2)/MSE
```

```
sumofCUMSIZE13<-sum(xdata$CUMSIZE*xdata$CUMSIZE3)/MSE
```

```
sumofCUMSIZE23<-sum(xdata$CUMSIZE2*xdata$CUMSIZE3)/MSE
```

```
sumofCUMINCCUMSIZE<- sum(xdata$CUMINC*xdata$CUMSIZE)/MSE
```

```
sumofCUMINCCUMSIZE2<- sum(xdata$CUMINC*xdata$CUMSIZE2)/MSE
```

```
sumofCUMINCCUMSIZE3<- sum(xdata$CUMINC*xdata$CUMSIZE3)/MSE
```

```
Q1<-c(sumofsquaresCUMSIZE,sumofCUMSIZE12,sumofCUMSIZE13)
```

```
Q2<-c(sumofCUMSIZE12,sumofsquaresCUMSIZE2,sumofCUMSIZE23)
```

```
Q3<-c(sumofCUMSIZE13,sumofCUMSIZE23,sumofsquaresCUMSIZE3)
```

```
QQ<-data.matrix(data.frame(Q1,Q2,Q3))
```

```
QQ
```

```
RR<- data.matrix(QQ+ggInverseMSE)
```

```
RR
```

Finding the estimate of the elements of \sum_{β}

```
RRInverse<- solve(RR)
```

```
RRInverse
```

Estimating the constant term \hat{C}

```
#cross product of a_sub i's and meanvector
```

```
CC<-MeanVector%*%ggInverse
```

```
CC
```

```
# calculate -0.5*C
```

```
WW<-
```

```
data.matrix(c(sumofCUMINCCUMSIZE,sumofCUMINCCUMSIZE2,sumofCUMINCCUMSIZE
3))
```

```
WW
```

```
EE<-WW+t(CC)
```

```
EE
```

Estimates of Bayesian regression parameters $\hat{\beta}$

```
FF<-RRInverse%*%EE
```

```
FF
```

Gini estimates

```
AA<- ((FF[1]*0.5)+(FF[2]/3)+(FF[3]*0.25))
```

```
AA
```

```
Gini<- (0.5-AA)/0.5
```

```
Gini
```

```
}
```

```
ginifxn(xdata)
```

Bootstrap estimation

```
ginibootstrap<- function(x,i){
```

```
  d<- x[i,] # allows boot to select sample
```

```
  return(ginifxn(d))
```

```
}
```

Bootstrapping with 1000 replications

```
results<- boot(data=xdata, ginibootstrap,
```

```
  R=1000)
```

View results

```
results
```

```
plot(results)
```

Appendix 3

Table 3A: Bayesian estimates for regression parameters (GLSS 5)

Geographical areas	Male headed households					Female headed households					All head of households			
	β_1	β_2	β_3	R-Square	Adj R-square	β_1	β_2	β_3	R-Square	Adj R-square	β_1	β_2	β_3	R-Square
Western	0.2819	0.1284	0.5080	0.999	0.998	0.2041	-0.0650	0.7052	0.998	0.993	0.2325	0.1741	0.4996	0.999
Central	0.2219	-0.0002	0.6257	0.997	0.996	0.3086	-0.6085	1.0850	0.992	0.987	0.2250	-0.1477	0.7510	0.997
Greater Accra	0.2971	-0.2400	0.8049	0.997	0.996	0.2181	-0.2610	0.8762	0.998	0.995	0.2338	-0.1114	0.7260	0.997
Volta	0.1684	0.2032	0.5387	0.999	0.998	0.2737	-0.1539	0.8020	0.999	0.994	0.1838	0.1478	0.5822	0.999
Eastern	0.1159	0.3440	0.4143	0.999	0.998	0.2213	-0.0867	0.7766	0.999	0.996	0.1255	0.2743	0.4841	0.998
Ashanti	0.2456	-0.1145	0.7535	0.998	0.998	0.2365	-0.2861	0.9355	0.998	0.996	0.2346	-0.1312	0.7800	0.998
Brong Ahafo	0.2245	-0.1795	0.7286	0.997	0.995	0.3165	-0.2859	0.8186	0.996	0.992	0.2430	-0.2091	0.7534	0.996
Northern	0.1823	-0.2171	0.8658	0.998	0.997	0.1261	0.1238	0.6551	0.996	0.985	0.1791	-0.1885	0.8465	0.998
Upper East	0.2840	-0.7243	1.3065	0.996	0.994	0.2386	-0.5289	1.1385	0.992	0.981	0.2857	-0.7201	1.2968	0.996
Upper West	0.2598	-0.2950	0.8903	0.998	0.995	0.4087	-0.9857	1.4445	0.992	0.977	0.2716	-0.3394	0.9278	0.997
Coastal	0.2231	-0.0561	0.6903	0.997	0.997	0.2379	-0.3378	0.9038	0.996	0.995	0.2007	-0.0749	0.7172	0.997
Forest	0.2018	0.1234	0.5559	0.999	0.999	0.2625	-0.2730	0.8981	0.999	0.998	0.2019	0.0610	0.6212	0.999
Savannah	0.2170	-0.4266	1.0283	0.998	0.997	0.1879	-0.2723	0.9739	0.998	0.996	0.2114	-0.3853	1.0001	0.997
Rural	0.1852	-0.1872	0.8577	0.999	0.998	0.2067	-0.1473	0.8073	0.998	0.998	0.1870	-0.1549	0.8262	0.999
Urban	0.2132	0.0314	0.6183	0.998	0.998	0.2364	-0.2749	0.8863	0.998	0.997	0.1951	0.0071	0.6567	0.998

Source: Author's computation from GLSS 5

Table 3B: Bayesian estimates for regression parameters (GLSS 6)

Geographical areas	Male headed households					Female headed households					All head of households			
	β_1	β_2	β_3	R-Square	Adj R-square	β_1	β_2	β_3	R-Square	Adj R-square	β_1	β_2	β_3	R-Square
Western	0.2795	-0.5083	1.0182	0.996	0.996	0.2386	-0.6569	1.2300	0.995	0.993	0.2543	-0.4915	1.0303	0.996
Central	0.2467	-0.7843	1.2013	0.988	0.987	0.2480	-1.0084	1.5583	0.994	0.992	0.2154	-0.7296	1.2044	0.983
Greater Accra	0.2321	-0.3769	0.9339	0.995	0.995	0.2138	-0.5359	1.0494	0.994	0.992	0.2150	-0.3804	0.9382	0.995
Volta	0.2517	-0.6051	1.0775	0.993	0.992	0.3151	-0.9367	1.4378	0.992	0.991	0.2423	-0.6047	1.0998	0.991
Eastern	0.2609	-0.3653	0.9225	0.996	0.996	0.2992	-0.7432	1.2785	0.994	0.993	0.2568	-0.4325	0.9989	0.996
Ashanti	0.3449	-1.0982	1.4254	0.983	0.982	0.2405	-0.7263	1.2030	0.989	0.987	0.3106	-0.9688	1.3354	0.980
Brong Ahafo	0.3054	-0.7185	1.0936	0.989	0.988	0.2706	-0.7171	1.0873	0.984	0.981	0.2742	-0.6623	1.0567	0.990
Northern	0.3509	-0.9699	1.3512	0.992	0.991	0.3587	-1.3843	1.8566	0.995	0.990	0.3457	-0.9775	1.3714	0.991
Upper East	0.2144	-0.5437	0.8931	0.991	0.991	0.3521	-1.0754	1.5128	0.986	0.983	0.2291	-0.6030	0.9670	0.990
Upper West	0.3670	-1.1811	1.4623	0.977	0.976	0.3253	-1.0617	1.4953	0.987	0.983	0.3410	-1.0838	1.3911	0.976
Coastal	0.2483	-0.5905	1.0870	0.993	0.992	0.2240	-0.7075	1.2383	0.994	0.993	0.2237	-0.5751	1.0954	0.992
Forest	0.2791	-0.6939	1.1286	0.992	0.992	0.2592	-0.7563	1.2559	0.993	0.993	0.2615	-0.6696	1.1295	0.992
Savannah	0.3007	-0.8848	1.2495	0.989	0.988	0.3433	-1.2004	1.5532	0.987	0.986	0.2928	-0.8777	1.2534	0.988
Rural	0.2691	-0.6647	1.1461	0.995	0.994	0.2838	-0.8388	1.3724	0.994	0.994	0.2652	-0.6652	1.1600	0.994
Urban	0.2600	-0.7347	1.1786	0.990	0.990	0.2302	-0.7512	1.2516	0.992	0.992	0.2391	-0.6986	1.1662	0.989

Source: Author's computation from GLSS 6

Table 3C: Bayesian estimates for regression parameters (GLSS 5 and GLSS 6)

Age group	Male headed households					Female headed households					All head of households			
	β_1	β_2	β_3	R-Square	Adj R-square	β_1	β_2	β_3	R-Square	Adj R-square	β_1	β_2	β_3	R-Square
5	0.16841	-0.1189	0.80329	0.998	0.998	0.2226	-0.2667	0.888	0.998	0.998	0.1772	-0.1311	0.8061	0.998
6	0.28053	-0.8401	1.26091	0.990	0.990	0.2617	-0.8871	1.3698	0.993	0.992	0.2683	-0.8145	1.2549	0.991

Source: Author's computation from GLSS 5 and GLSS 6