Joint Modelling of Child Poverty and Malnutrition in Children Aged 6 to 59 Months in Malawi

UNIVERSITY OF KWAZULU - NATAL
INYUVESI YAKWAZULU-NATALI

Lindani Dube

December, 2019

Joint Modelling of Child Poverty and Malnutrition in Children Aged 6 to 59 Months in Malawi

by

Lindani Dube

A thesis submitted to the
University of KwaZulu-Natal
in fulfilment of the requirements for the degree
of
MASTER OF SCIENCE
in
STATISTICS
under the supervision of

Ms. Danielle Jade Roberts

UNIVERSITY OF KWAZULU-NATAL
SCHOOL OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE
WESTVILLE CAMPUS, DURBAN, SOUTH AFRICA

UNIVERSITY OF KWAZULU-NATAL

SCHOOL OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE

WESTVILLE CAMPUS, DURBAN, SOUTH AFRICA

Declaration - Plagiarism

I, Lindani Dube, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowlegded as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then
   (a) their words have been re-written but the general information attributed to them has been referenced, or
   (b) where their exact words have been used, then their writing has been placed in italics and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the reference sections.

Lindani Dube (Student)	Date
Danielle Jade Roberts (Supervisor)	Date

Disclaimer

This document describes work undertaken as a Masters programme of study at the University of KwaZulu-Natal (UKZN). All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institution.

Abstract

The objective of this study was to identify risk factors associated with poverty and malnutrition of children among the ages 6-to-59 months in the country of Malawi, making use of the joint model. By joint modelling, we refer to simultaneously analysing two or more response variables emanating from the same individual. Using the 2015/2016 Malawi Demographic and Health Survey, we jointly examine the relationship that exists between poverty and malnutrition of children among 6-to-59 months in Malawi. Jointly modelling these two outcome variables is appropriate since it is expected that people that live under poverty would have a poor nutrition system, and if a child is malnourished, the likelihood that they come from a poor family is greatly enhanced. Jointly modelling correlated outcomes can improve the efficiency of parameter estimates compared to fitting separate models for each outcome, as joint models have better control over type I error rates in multiple tests. A generalized linear mixed model (GLMM) was adopted and a Bayesian approach was used for parameter estimation. The potential risk factors considered in this study comprised of the childs age in months, gender of child, birth weight, birth order, mothers education level, head of household sex, language, household smoking habit, anaemic level, type of residence (urban or rural), region, toilet facility, source of drinking water, and multiple births. Each response was modelled separately as well as jointly and the results compared. The R package *MCMCglmm* was used in the analyses. The joint model revealed a positive association between malnutrition of children and poverty in the household.

Acknowledgements

This work was sponsored by the Moses Kotane Institute. The Moses Kotane Institute is an entity of the KwaZulu-Natal provincial government through the Department of Economic Development, Tourism and Environmental Affairs (EDTEA). Its mandate is to provide an internationally recognised quality research driving the economic development and innovation in the province. My deepest gratitude is extended to this institute for sponsoring my studies. The views expressed in this publication are those of the author(s) and not necessarily those of the Moses Kotane Institute or the EDTEA.

It is my desire to express deep and sincere gratitude to my supervisor, Ms Danielle Jade Roberts whose encouragement, diligence, expertise, understanding, patience and professionalism added tremendous quality to my Master's research project. I appreciate all her contributions of time and ideas, which made my thesis experience productive and stimulating. The passion and enthusiasm she has for her research was contagious and inspirational to me, even during tough times in my studies. I would like to extend many thanks to Ms Nombuso Zondo for taking her precious time to take me through the Statistical Analytic Software (SAS).

Finally, wonderful appreciation goes to my family. I appreciate all the love, support and guidance they have given me throughout my whole entire life. If it was not for their beautiful souls, I would not have made it this far.

Contents

List of Figures

List of Tables

Abbreviations

AR(1)	Autoregression order of (1)
CRLB	Cramer Rao Lower Bound
CS	Compound Symmetry
DHS	Demographic and Heath Survey
DIC	Deviance Information Criteria
GLM	Generalized Linear Models
GLMM	Generalized Linear Mixed Models
GLSE	Generalized Least Squares Estimate
GS	Gibbs Sampler
LM	Linear Model
LMM	Linear Mixed Model
MDHS	Malawi Demographic and Heath Survey
MCMC	Makov chain Monte Carlo
ML	Maximum Likelihood
M-H	Metropolis-Hastings
NSO	National Statistical Office
OLSE	Ordinary Least Squares Estimate
OR	Odds Ratio
SAS	Statistical Analytic Software
SEA	Standard Enumeration Area
Teop	Toep and Toeplitz
UN	Unstructured
USAID	United States Agency for International Development

Chapter 1

Introduction

1.1 Background on Poverty and Malnutrition

Poverty, in short, is defined as the state of not having enough to meet the basic needs (Investopedia, 2018), and malnutrition is defined as a condition that arises from consuming a diet in which nutrients are either insufficient or are too much such that the diet causes health problems (Richmond Vale Academy, 2016). This may involve calories, protein, carbohydrates, vitamins or minerals (Blössner et al., 2005). Not enough nutrients is referred to as undernutrition or undernourishment, while too much is referred to as overnutrition. This section serves as the motivation for this study, and it examines some of the important detrimental effects of poverty and malnutrition on children.

Poverty is a predicament of being extremely poor, the state of being inferior in quality or insufficient in amount (Investopedia, 2018). There are various types of poverty (eschool Today, 2008).

- Absolute: The extreme kind of poverty that involves the chronic lack of basic needs.
- Relative: Usually in relation to other members and families in the society.
- Situational (transitory): Normally due to some adversities like natural disasters or serious illness.
- Generational or chronic poverty: poverty that is handed over to individuals and families from generations before them.

The threshold at which absolute poverty is defined is considered to be about the same, independent of the person's permanent location or era (Addae-Korankye, 2014). The threshold at which relative poverty is defined, varies from one country to another, or from one society to another. Most developing countries tend to be more vulnerable to poverty due to their poor infrastructure and very low economic growth.

Eradicating poverty was a part of the first target of the Millennium Development Goals (eradicate extreme hunger and poverty), and is now goal 1 of the Sustainable Development Goals (end poverty in all its forms everywhere). There has been considerable progress in reducing poverty over the past decades (The World Bank, 2018). However, despite this progress, the number of people living under extreme poverty globally still remains significantly high. Further, more than half of the extreme poor live in sub-Saharan Africa. In fact, this region has seen a rise in poverty rates from 1990 to 2015 (The World Bank, 2018). Malawi, a sub-Saharan African country, is considered one of the poorest countries in the world, with three-fifths of its population being unable to meet their basic daily requirements (Davids & Gouws, 2013; The World Bank, 2018). Malawi has a very young and rapidly growing population, which is a key factor of Malawis persistent poverty, where efforts to reduce poverty have failed (International Monetary Fund, 2017).

Children living in poverty are the most impacted as it has immediate and lasting consequences that can follow a child into adulthood. Research shows that children from families that live under poverty tend to have poor vocabulary, poor communication skills, lack ability to concentrate and struggle with numbers (Birdsong, 2016; Project, 2014). Consequently, due to the long term effects of poverty, this in turn results in a country's poor level of education. Van der Berg (2008) showed an interrelationship between poverty and education. While poor people do not have enough to access quality education, a lack of adequate education also results in poverty. Moreover, the fact that poor people are provided with weak quality education condemns them to low-wage employment or even unemployment, thus hinders their human potential being unlocked to allow them to succeed and thrive (Project Rise, 2018). Poverty results in very many stressors such as family conflict, economic strain, exposure to discrimination, transitions and many more other traumatic events that can negatively impact on the behaviour of children living in such conditions (Greever, 2014).

People struck with poverty are more likely to engage in criminal activities due to their inability to secure steady or sufficient financial resources, which often leads to illicit activities to generate income (Oxford Handbooks Online, 2017). It is true that not all criminals are poor and not all poor are criminals, however the chance of people living under poverty to commit crimes brings about the correlation. History suggests a direct connection between poverty and crime in reality. In the essay done by Essay Professors (2006), it was argued that people often engage in criminal activities primarily because of fear of poverty; while people living under poverty will do anything to escape the circle of poverty, and the rich also will do everything to remain rich. Furthermore, poverty results in hopelessness, lack of motivation, depression and anger, which can lead people into crimes without considering the consequences (Essay Professors, 2006). Thus, efforts must be made to eradicate poverty in order to bring about radical change.

Poor individuals also tend to have a shorter life expectancy compared to rich individuals (Poverties, 2011). In the study that was done by the Health Central (2016) in the United States, it was discovered that people located in the poorest areas died 10 years earlier than those located in the wealthiest areas. Living under poverty suppresses means to afford or to have access to proper and decent health care, and for this reason, poor individuals can be condemned to die younger (World Health Organization, 2005; The World Bank, 2014). Many researchers have agreed that education and health are fundamental to poverty reduction (World Health Organization, 2005). Thus, it goes without saying that poverty means poor living conditions, poor diet, poor education, and poor medical assistance, and therefore it becomes a challenge to escape the burden of poverty (UK Essays, 2015). When looking at the socio-economic status, both at micro (individual) and macro (society) levels, important links can be observed between poverty and health status, particularly between poverty and malnutrition (Phipps, 2003).

Malnutrition can cause serious health problems, both for children and adults (Nursing Times, 2009). Consuming an excess amount of calories and nutrients usually results in overweight and obesity, which may lead to serious medical conditions, such as Type 2 diabetes mellitus, coronary heart disease, stroke, and high blood pressure, among others (Live Strong, 2017). The fact that the effects of many of these chronic diseases cannot be reversed, implies that malnutrition in children must be tackled before they reach the age of two years, or else it may result in serious problems in the long run (Richmond Vale Academy, 2016). Typical symptoms of malnutrition in children includes dizziness, scaly and dry skin, underweight, poor growth, muscle weakness, learning difficulties, decaying teeth, bloated stomach, swollen and bleeding gums, problems with organ function, fatique and low energy, weak bones, poor immune function (Weiss, 2016; Nursing Times, 2009). Moreover, malnutrition can inhibit a child's ability to achieve the normal height or body weight (Live Strong, 2017). In developing countries, the most common case is usually that of undernutrition, when a person is not getting enough calories, protein, or micronutrients, or due to not enough high-quality food being available to eat as a result of high food prices and poverty (World Health Organization, 2018b). Webb & Bhatia (2005) found that if undernutrition occurs during pregnancy, or before two years of age, it may result in permanent problems with physical and mental development. Efforts to improve nutrition are some of the most effective forms of development aid.

In recent decades, malnutrition of children has emerged as a serious challenge on the global development agenda (Rajkumar et al., 2011). Malnourished children are said to have a decreased ability to focus on tasks, an increased emotionality, listlessness, apathy, a lack of interest in the environment and consequently, this will certainly have a negative impact on their long-term intellectual development (Amcoff, 1981). According to World Health Organization (2018a), in 2018 the prevalence of stunting in children globally was 22% with 17 million young children affected by wasting in its severe form. Stunting and wasting are indicators of malnutrition, specifically undernutrition, in children. Children that are undernourished are exposed to a higher risk of chronic diseases related to nutrition. Research shows that more than 25% of children younger than the age of 15 years in sub-Saharan Africa are underweight due to poor nutrition intake, of which increases their chances of being vulnerable to diseases and less productive in their school work (Chinyoka, 2014). Thus, if a child has a poor nutritional system history, the likelihood that they battle with their school work is greatly enhanced (Jamison, 1986).

While the effects of poverty on malnutrition in children is apparent, high rates of malnutrition in children can also affect the future economic growth of a country as it reduces the intellectual, physical and emotional potential of the population (Marini & Gragnolati, 2003). Thus, malnutrition also contribute to poverty.

1.2 Literature Review

The interrelationship between poverty and malnutrition in young children has been demonstrated in previous studies (Habyarimana et al., 2016; Yimer, 2000). These studies also revealed a positive correlation between the responses. Habyarimana et al. (2016) revealed that the child's age, birth order, gender, birth weight, multiple birth of the child, fever, anaemia of the mother, body mass index of the mother, mother's education level and knowledge on nutrition, age of household head, source of drinking water, toilet facilities, place of residence of household and region were jointly significantly associated with poverty and malnutrition in children aged 6 to 59 months old in Rwanda. A study by Hunt (2005) discusses the potential impact of reducing global malnutrition on poverty reduction and economic development, where the author also discusses evidence that suggests that thin children are usually from thin mothers, such that both the mother and child get caught up in an intergenerational cycle of poverty and malnutrition.

In addition, many studies have linked poverty and malnutrition in children to a number of factors, each separately. In the study that was done in Dollo Ado District, Ethopia, it was discovered that the child's age and gender, mother's education level and marital status, and having livestock were among the key factors associated with malnutrition in children aged 6 to 59 months. The study further suggested a significant impact of household income on the likelihood of malnutrition in a child (Demissie & Worku, 2013). Similarly, a study that was conducted by Yimer (2000) affirmed that the household's economic status, mother's education level, child's age, preceding birth interval and the number of under-five children were important in explaining the variation in long-term nutritional status of children.

Ajakaiye & Adeyeye (2001) argued that the age and education level of different household members (head), number of income earners, household composition and size, assets owned by household, access to basic social services, sex and ethnicity of household head, location variable (rural or urban), sector of employment and remittances to household were among the significant factors associated with poverty. A positive correlation was found between earnings and higher education in the study that was done in Malaysia by Mok et al. (2007). Likewise, the study further revealed that household size, race and regions were crucial factors linked to poverty in urban Malaysia. In addition, Omole (2003) found that loss of employment, low income of family head, low income by other family members, lack of access to farm land, protracted sickness of family head, low educational qualifications of family head, low educational qualifications of other family members, death of family head/breadwinner, loss or death of livestock, neglect by government, poor economic/business condition were significant risk factors of poverty.

With regard to poverty in Malawi, poor households are generally larger than more wealthier households. In the study that was done by Wood & Mayer (2006), it was found that when looking at the average household size by income decile, households in the poorest decile were more than twice as large as households in the richest decile. The larger household size is usually reflective of the higher number of dependents in the poorer households. An older study done on the determinants of poverty in Malawi, showed that higher levels of educational attainment, especially in women, and the reallocation of household labour away from agriculture and into the trade and services sector can be effective in reducing poverty in Malawi (Mukherjee & Benson, 2003).

1.3 Aims and Objectives

This study assesses the prevalence of children living in poverty as well as the prevalence of malnutrition in children aged 6 to 59 months in Malawi. In addition, this study aimed at achieving the following objectives:

- To determine the significant socio-economic and demographic risk factors associated with child poverty in Malawi.
- To determine the significant socio-economic and demographic risk factors associated with malnutrition in children aged 6 to 59 months in Malawi.
- To examine the correlation between poverty and malnutrition in the children.
- To investigate the joint effects of socio-economic and demographic risk factors on child poverty and malnutrition in the children.
- To compare the results of the separate models to that of the joint model.

1.4 Thesis Structure

Chapter 2 introduces the data that will be used in the analysis and discusses the statistical models appropriate for the study. Chapter 3 gives a brief overview of the generalized linear model and extends it to the logistic regression model. Chapter 4 discusses the generalized linear mixed model and introduces the Bayesian approach to estimation of that model. In addition, Chapter 4 gives an overview of the generalized linear mixed model approach to joint modelling of multiple responses. The results of the analysis are discussed in Chapter 5 and Chapter 6 discusses the findings and conclusions of the study.

Chapter 2

Data Description and Exploratory Data Analysis

2.1 Study Area

The Republic of Malawi, that was formerly known as Nyasaland, is a landlocked country that is located in Southern Africa. Malawi is bordered by Tanzania to the north east, Zambia to the north west and Mozambique to the south, east and west. Malawi is separated from Tanzania and Mozambique by the third largest lake in Africa called Lake Malawi and covers about one fifth of the countries total area (Cross-Border Road Transport Agency, 2016). Figure 2.1 on page 8 shows the map of the country of Malawi.

Malawi has a land area of over 45 560 square miles with its population estimated around 19 050 460 in 2018. Malawi is considered among the world's least developed countries with its economy centered on agriculture and a population that is mostly rural with a Human Development Index of 0,418 according to United Nations. The government of Malawi deals with problems in improving the economy, education, healthcare, environmental protection, and becoming financially independent amidst widespread overpopulation and unemployment. Thus, Wu House (1992) concluded that rapid growth of population is exerting extreme pressures on land, employment, education and health and therefore inhibits the nation's ability to satisfy the most basic of human needs.

The economy of Malawi primarily relies on agriculture with approximately 90% of the whole population residing in the rural areas. Substantial inflow of economic assistance from the International Monetary Fund (IMF), the World Bank, and individual donor nations plays a significant role in assisting Malawi's economy.

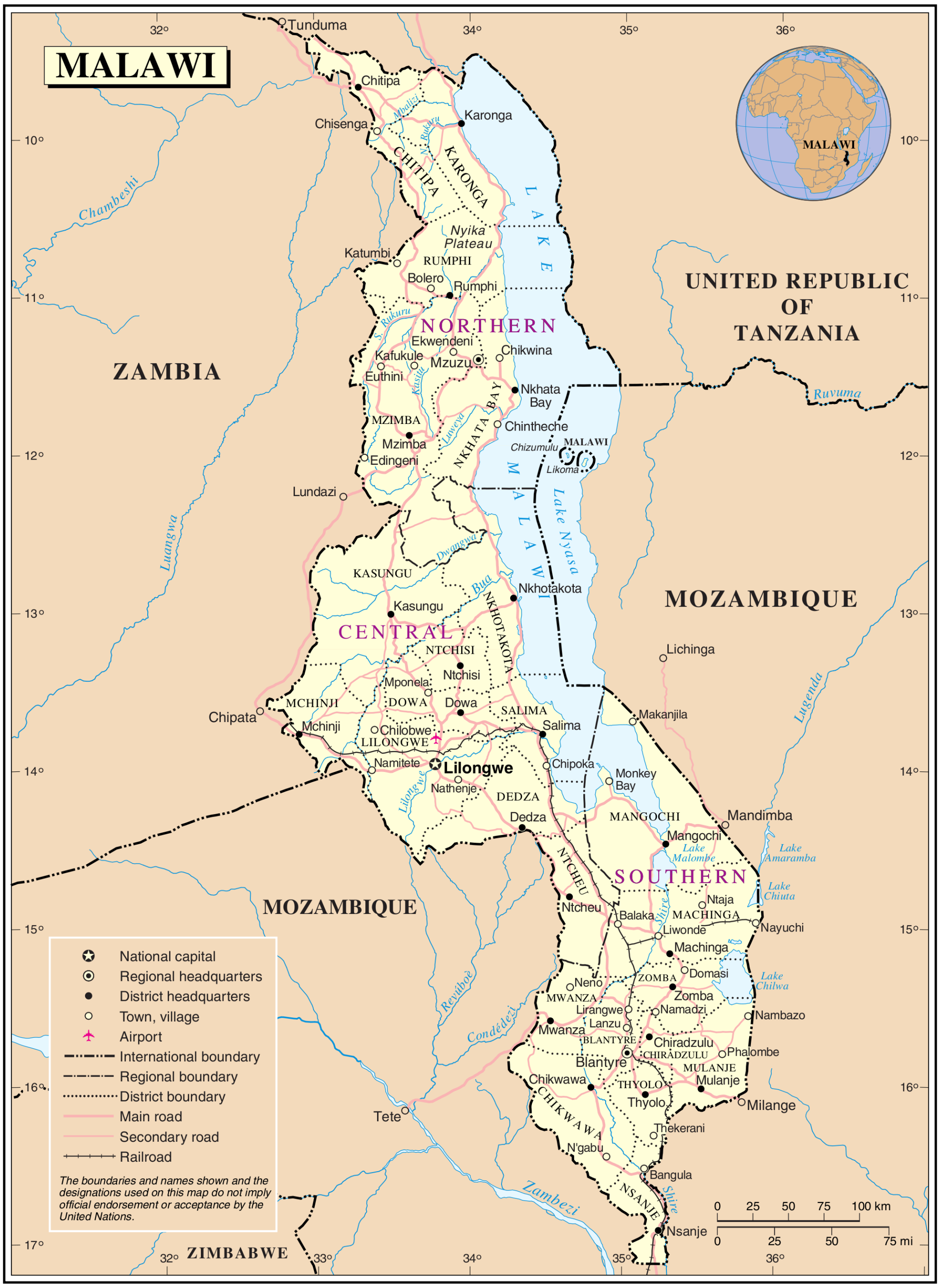


Figure 2.1: The map of Malawi

Tobacco is one of the highest productions in Malawi. However, this also creates a burden on the country's economy as world tobacco prices decline and the international community increases pressure to limit the production of tobacco. Moreover, the country also strongly depends on the production of tea, sugarcane, coffee, cotton, corn, potatoes, sorghum, cattle and goats, where agriculture accounts for about one third of the country's gross domestic product (GDP) and 90% of export revenues

(Cross-Border Road Transport Agency, 2016).

2.2 The Data

The data set utilized in this study is the 2015-16 Malawi Demographic and Health Survey (MDHS) that was collected from 19 October 2015 to 17 February 2016. The program aimed to collect and analyse reliable demographic indicators and health estimates at regional and family levels in order to eventually provide better health interventions. The whole purpose of the 2015-16 MDHS was to assist the country of Malawi to come up with strategies, evaluate and design programmes to improve the health of the countrys population.

2.2.1 Sampling Design and Data Collection

The MDHS was nationally represented and utilised a stratified multi-stage cluster design. The Malawi population was first divided into 28 districts, following this the population was then classified into categories according to region (rural or urban), which yielded 56 sampling strata. The two strata were then sampled independently in two stages. The first stage consisted of selecting standard enumeration areas (SEAs) or clusters proportional to their size. The second stage was made up of a fixed number of 30 households per urban cluster and 33 per rural cluster that were selected with an equal probability systematic selection. Thereafter, all men (15-54 years old) and women (15-49 years old) that were either permanent residents of the selected households or visitors who stayed in the household the night before the survey were eligible to be interviewed.

Four questionnaires were used in the 2015-16 MDHS: the Household Questionnaire, Woman's Questionnaire, Man's Questionnaire, and Biomarker Questionnaire. These questionnaires were designed to collect information regarding the characteristics of the household and eligible women, men and children. The Household Questionnaire collected basic information on the characteristics of each member and recent visitors of the household, including age, sex, and relationship to the head of the household. This questionnaire also collected information on characteristics of the household's dwelling unit, such as source of water; type of toilet facilities; materials used for the floor, roof and walls of the house; and ownership of various durable goods. The Women and Men's Questionnaires were used to collect a range of information from all eligible women and men in the selected households. In a subsample of households selected, the 2015-16 MDHS incorporated the biomarkers for anthropometry, anaemia testing, and HIV testing. All children aged 6 to 59 months in these selected households were tested for anaemia with the parents or guardians consent. Furthermore, the weight and height measurements were obtained from eligible children aged 0-59 months (National Statistical Office (NSO) [Malawi] and ICF, 2017).

2.3 Study Variables

2.3.1 Dependent Variables

Listed below are the two dependent or response variables of interest in this study, namely malnutrition in children and child poverty.

Malnutrition

There are three anthropometric indicators that can be used for monitoring malnutrition in children: stunting (height-for-age); under weight (low weight-for-age); and wasting (low weight-for-height). These indicators represent children whose height-for-age, weight-for-age, weight-for-height fall more than two standard deviations below the median of internationally accepted growth standards. In this study, only height-for-age as an indicator of malnutrition was considered as it is a result of long-term nutritional deprivations (World Health Organization, 2017). Hence, malnutrition was based on whether or not a child was stunted, which was a binary outcome; yes or no. Children are defined as stunted if their height-for-age is more than two standard deviations below the Child Growth Standards median (i.e. if their Z-score is less than -2), and thus the calculation for stunting in this study was based on the following (World Health Organization, 2006):

$$\text{Z-score} = \frac{(\text{observed height}) - (\text{median reference height})}{\text{standard deviation of reference population}}$$

Thus, the higher the value of the Z-score, the better the nutritional status of the child.

Poverty

Child poverty refers to a child living under poverty, which is based on the household's poverty status. In this study, the MDHS household wealth index Z-score was used in classifying a household's poverty status. A household wealth index is a composite measure of a household's cumulative living standard and is calculated based on the ownership of various household assets (Croft et al., 2018). Each household was assigned a standardised score for each asset, where the score differed depending on whether or not the household owned that asset. These scores were then summed for each household. In this study, the households were ranked according to their wealth index Z-score and then divided into quintiles, that being five groups with an equal number of households in each. This resulted in a household wealth quintile. Children residing in households in the two poorest quintiles were regarded as living under poverty, which is a relative measure of poverty. Thus, this response variable was also binary, indicating whether or not the child was living under poverty.

2.3.2 Independent Variables

The independent variables of interest comprised of a range of socio-economic and demographic factors.

Malnutrition

The potential risk factors of stunting, our measure of malnutrition in this study, are displayed in Figure 2.2 below. These factors were either at individual level or household level.

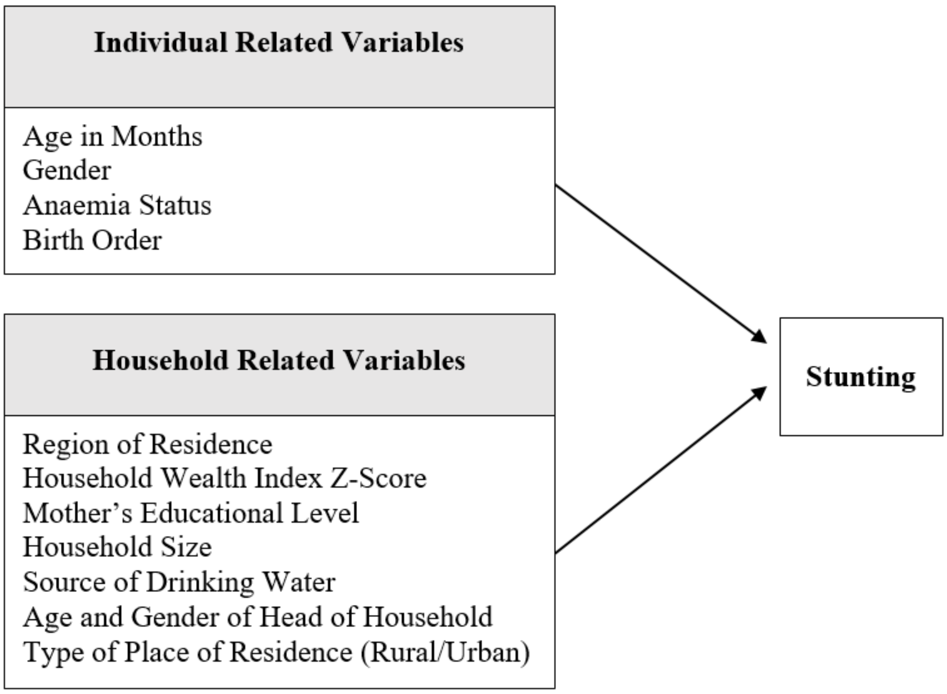


Figure 2.2: Potential risk factors of stunting in children

Poverty

The potential risk factors of child poverty are based on those factors at a household level, shown in Figure 2.3 below. These household factors considered are the same as those considered for malnutrition in children, excluding the household wealth index Z-score, as that is what poverty in households was based on. Although other studies have considered the education level of the head of the household as a risk factor of poverty, data on this variable was not available. Therefore, the mother's education level was used instead.

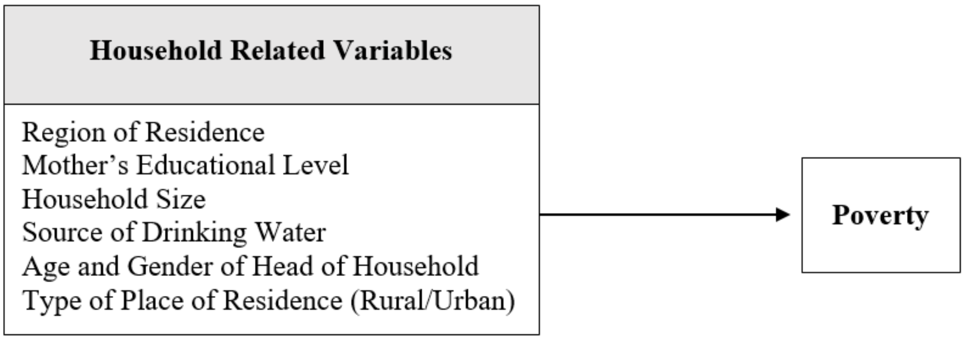


Figure 2.3: Potential risk factors of poverty in households

2.4 Exploratory Data Analysis

This section provides a better understanding of any sort of relationship that exists between the response variables and independent variables of interest. The final data set for this study was made up of 5123 children from 4150 households.

Table 2.1 displays some of the characteristics of this sample, which consisted of an almost equal number of male and female children. The majority of the children resided in households headed by males (72.1%) and resided in rural areas (84.5%). In addition, the majority of children had mothers with either no education (11.7%) or only a primary level of education (59.5%). More than half of the children resided in households with unprotected drinking water sources (63.6%), which comprised of open wells (private and public), unprotected springs, rainwater and surface water (rivers/streams, ponds/lakes and dams). Protected drinking water sources included protected wells (private and public), boreholes and protected springs, of which only 14.5% of children resided in households with access to this source of drinking water. Furthermore, 21.8% of the children resided in households with access to tap water for drinking, which included standpipes or public taps, water piped into the household or water piped into the yard. The overall observed prevalence of stunting in children was 30.5% and the overall observed prevalence of children living under poverty was 40.9%.

The uncorrected correlation between child poverty and stunting was estimated at 0.11 and was statistically significant at a 5% significance level.

Table 2.1: Characteristics of the sample

Variable	Sample size (%)	Variable	Sample size (%)
Gender		**Type of residence**	
Male	2532 (49.4)	Urban	795 (15.5)
Female	2591 (50.6)	Rural	4328 (84.5)
Age in months		**Region**	
$6-11$	541 (10.6)	Central region	946 (18.5)
$12-23$	1095 (21.4)	Northern region	1796 (35.1)
$24-35$	1148 (22.4)	Southern region	2381 (46.5)
$36-47$	1193 (23.3)	**Mother's education level**	
$48-59$	1146 (22.4)	No education	597 (11.7)
Anaemic level		Primary	3050 (59.5)
Severe	88 (1.7)	Secondary or higher	1037 (20.2)
Mild	1721 (33.6)	Don't know	439 (8.6)
Moderate	1396 (27.2)	**Source of drinking water**	
Not anaemic	1918 (37.4)	Protected water	745 (14.5)
Stunted		Unprotected water	3259 (63.6)
Yes	1561 (30.5)	Tap water	1119 (21.8)
No	3562 (69.5)	**Head of household sex**	
Living under poverty		Male	3692 (72.1)
Yes	2097 (40.9)	Female	1431 (27.9)
No	3026 (59.1)		

From Figure 2.4, we can observe that stunting was more prevalent in children residing in the poorest households. This figure also indicates that there was a clear decline in the prevalence of stunting in children as the household wealth increased, therefore suggesting that poverty may be a fairly strong driver of malnutrition in children.

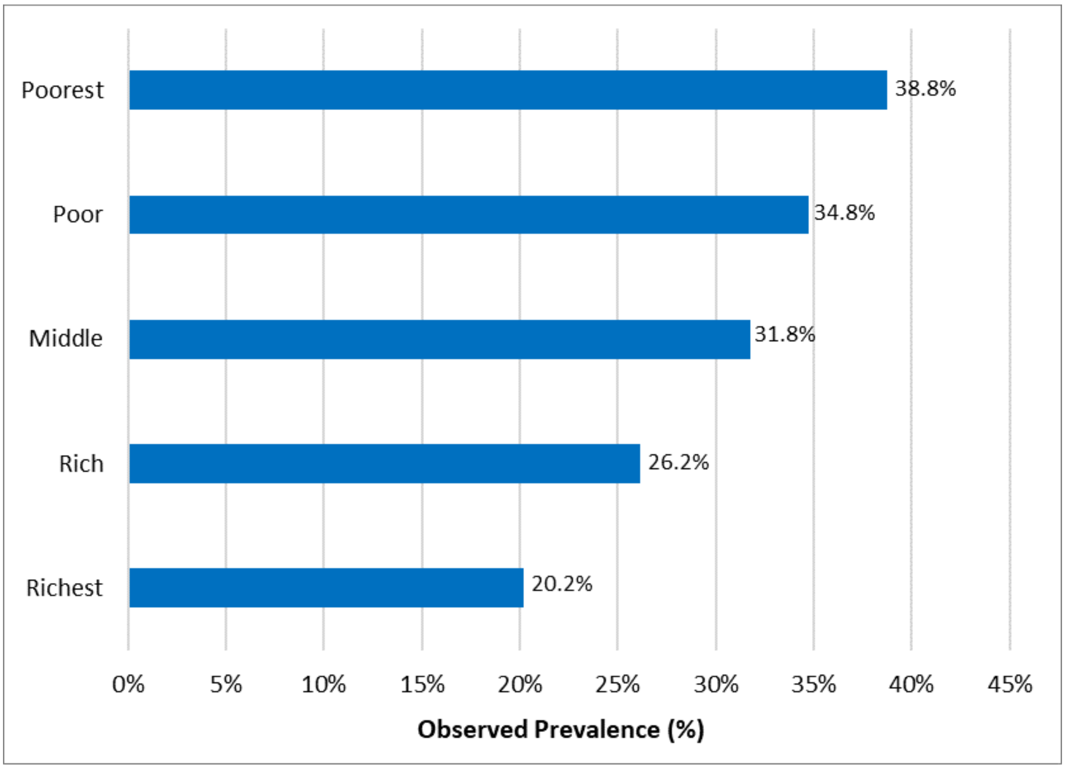


Figure 2.4: Observed prevalence of stunting according to household wealth quintile

Figure 2.5 on the next page displays the observed prevalence of stunting in children and child poverty according to the region of Malawi. The Central region had the highest prevalence of stunting and child poverty at 32.3% and 48.2%, respectively. The lowest prevalence of each was found in the Northern region.

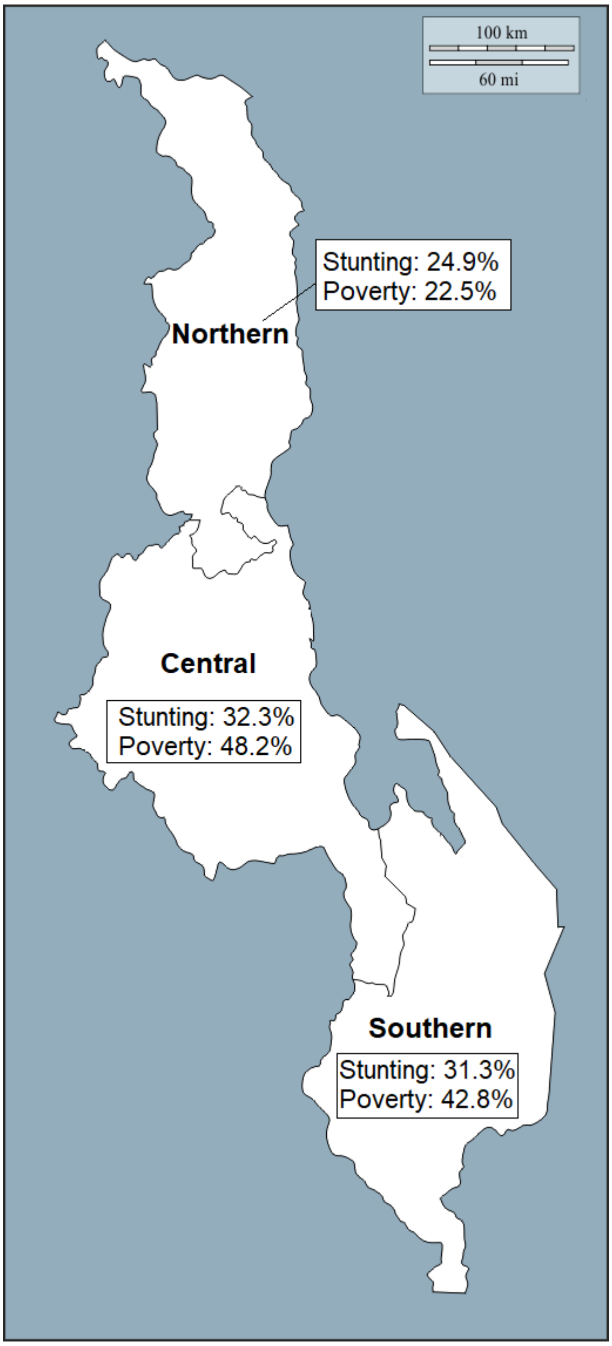


Figure 2.5: Observed prevalence of child poverty and stunting according to the region of Malawi

From Figure 2.6 on the next page, we can observe that the prevalence of both stunting and poverty was almost equal between male and female children, therefore suggesting that no significant gender bias was evident. This is consistent with the patterns observed in other African countries and with previous findings from Malawi.

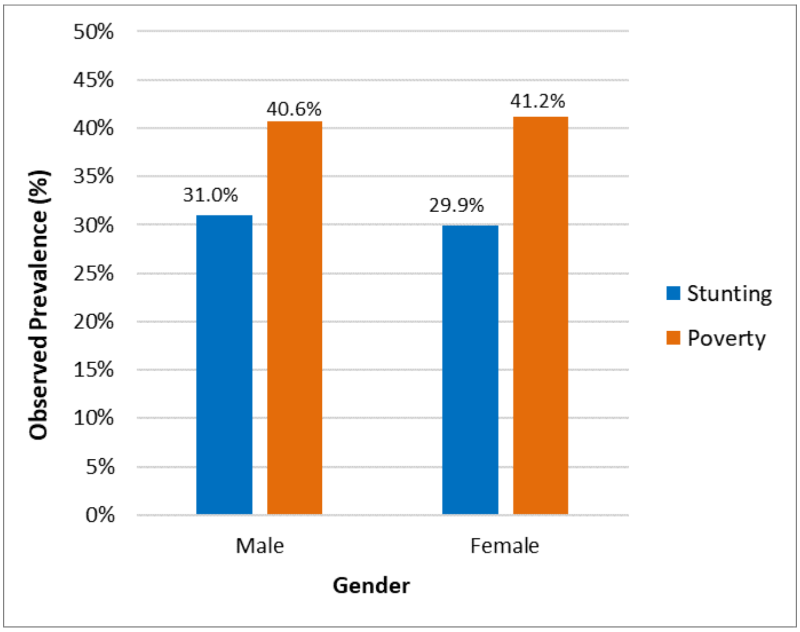


Figure 2.6: Observed prevalence of child poverty and stunting according to gender

Figure 2.7 on the next page shows that the lowest prevalence of stunting was in the age group of 6 to 11 months, after which the prevalence significantly increased. According to Mzumara et al. (2018), early childhood development is vital during the first 2.5 years of life, after which the effects of stunting is irreversible. In addition, malnutrition is considered a cumulative phenomenon (Marini & Gragnolati, 2003). This can be seen in Figure 2.7 with the older age groups having higher prevalences. Even though there was a decline in the prevalence of children living under poverty with an increase in age, the range of the prevalence across the age groups was low (between 38.1% and 42.5%).

Based on Figure 2.8 on the next page, an association between the child's anaemia status and both stunting and poverty can be observed, where there was a decline in the prevalence of both as the anaemia status improved. This is unsurprising as anaemia can be caused by nutritional deficiencies, which is also a cause of stunting. In addition, those who cannot afford foods rich in nutrients are prone to nutritional deficiencies. In the case where anaemia is severe, the prevalence of stunting and poverty are higher than where the anaemia categories are mild, moderate and not anaemic.

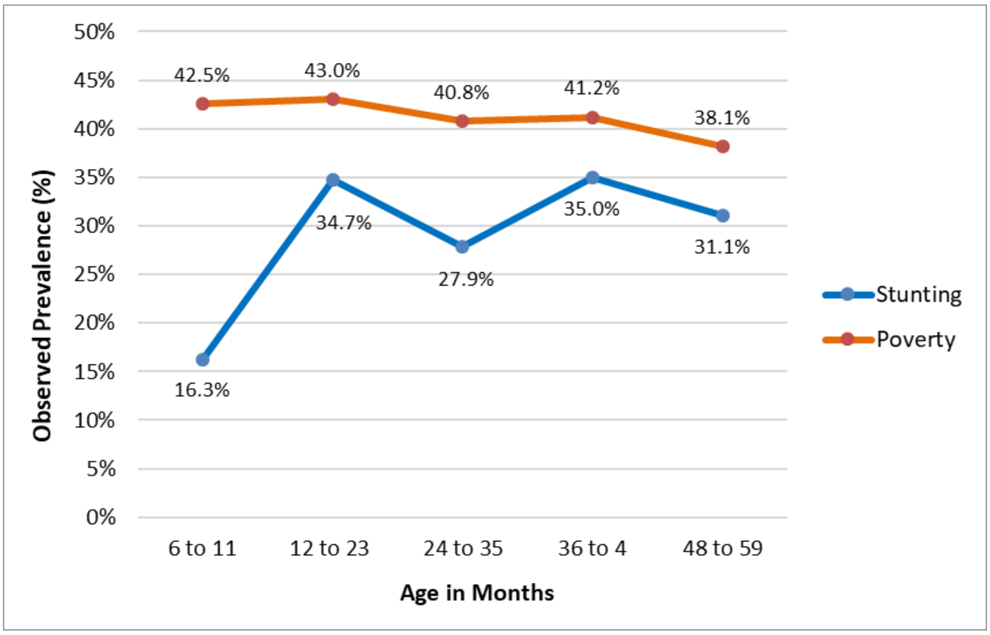


Figure 2.7: Observed prevalence of child poverty and stunting according to age in months

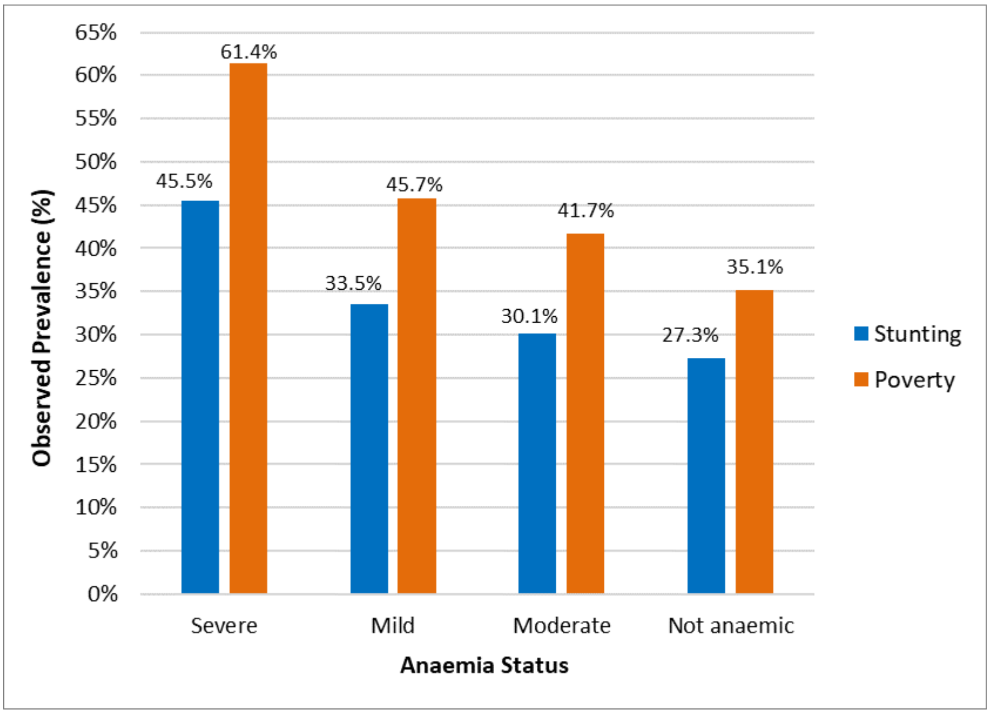


Figure 2.8: Observed prevalence of child poverty and stunting according to anaemia status

Figure 2.9 displays the prevalence of stunting and child poverty according to the birth order of the child. Not much variation is visible in the prevalences. Thus, there may not be a significant association between birth order and the two responses.

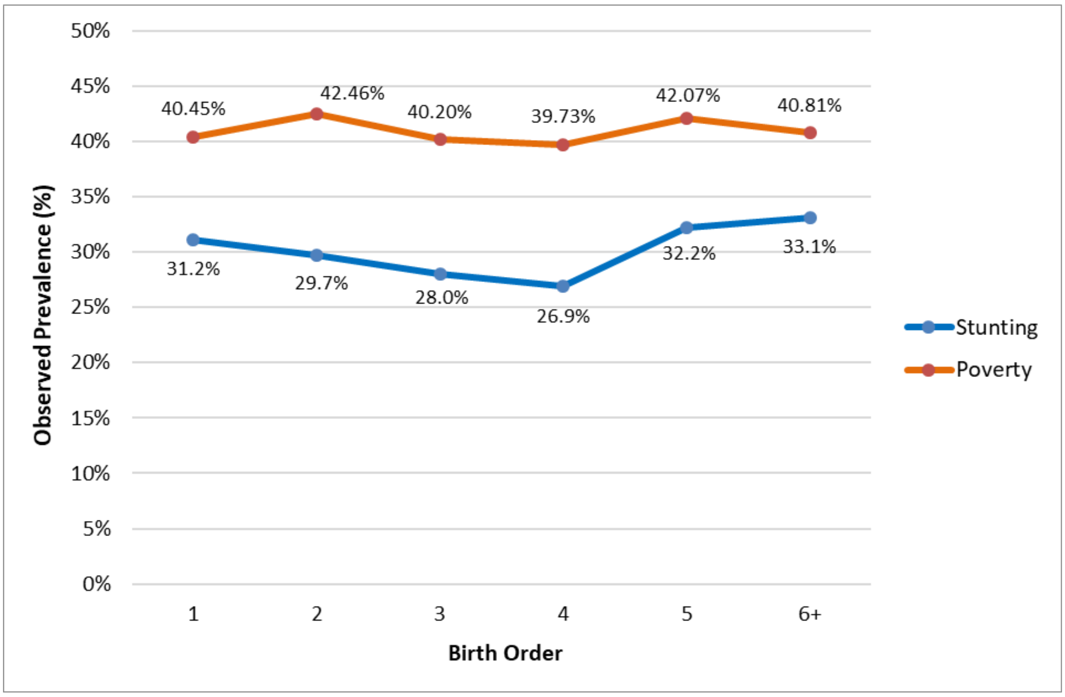


Figure 2.9: Observed prevalence of child poverty and stunting according to birth order

Figure 2.10 on the next page reflects an increase in the prevalence of stunting as the number of members in the household increased to 5. In contrast, the prevalence of child poverty decreased with an increase in the number of members in the household. This may be attributed to the fact that more household members means that there is more of a contribution to the household's wealth.

Poverty was most prevalent amongst children residing in households headed by younger individuals (see Figure 2.11 on the next page). This may be as a result of a lower earning potential of younger individuals as they do not yet have the experience and/or qualifications required for higher paying jobs. The prevalence of stunting displayed little variation in the lower age groups of the household head, however the prevalence varied significantly after the 46 to 55 age group. This may be due to small sample sizes in these age groups with only 10.6% of the children residing in households headed by individuals 56 years and older.

Figure 2.10: Observed prevalence of child poverty and stunting according to the household size

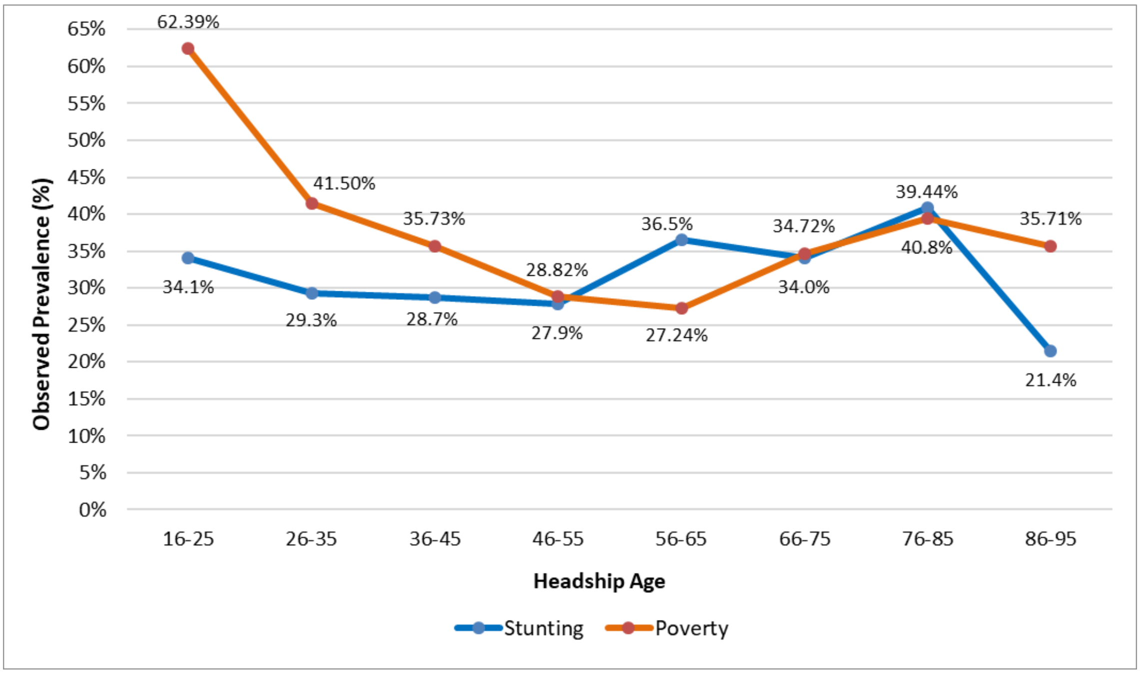


Figure 2.11: Observed prevalence of child poverty and stunting according to the age of the household head

The prevalence of child poverty and stunting in children for household characteristics, source of drinking water and type of place of residence, is shown in Figure 2.12 below. The prevalence of both was lowest amongst children residing in households with access to tap water as well as those residing in households in urban areas. As one would expect, the majority of children living in households with an unprotected drinking water source were living under poverty (59.3%), these children also had the highest prevalence of stunting (34.8%).

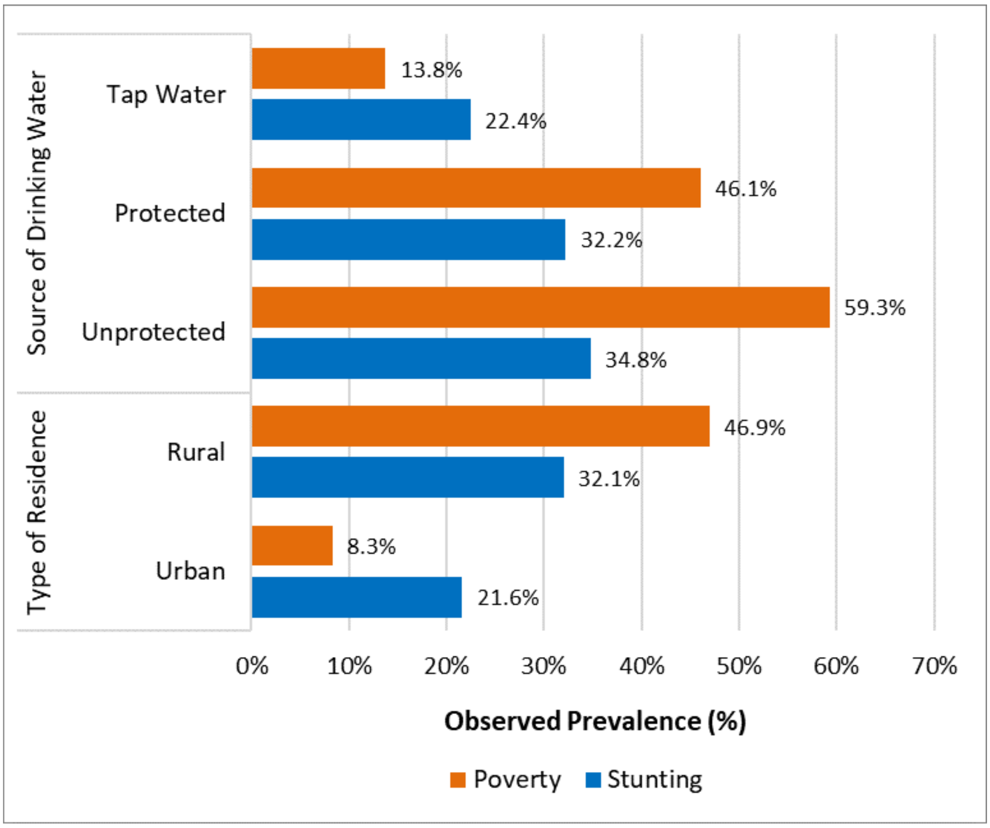


Figure 2.12: Observed prevalence of child poverty and stunting according to source of drinking water and type of place of residence

Figure 2.13 on the next page shows that the prevalence of stunting was similar amongst children who resided in households headed by males and females. However, there was a substantial difference in the prevalence of poverty amongst these children, with 50.7% of children living under poverty residing in households headed by females. The data suggested that the mother's highest education level played a crucial role in the prevalence of malnutrition in children aged 6 to 59 months. The observed prevalence of malnutrition, as measured by stunting, was much lower among children whose mother had more than a primary education (Figure 2.13). Figure 2.13 shows vividly that the higher the mother's educational level, the lower the prevalence of stunting. Although we do not know any information regarding the 'unknown' category of the mother's education level, with a second highest prevalence of stunted children, we can assert that these mothers could be poorly educated. The highest prevalence of children living under poverty was observed for those with mothers who have no education, followed by those with mothers who have only a primary education. This is unsurprising considering education has a significant effect on job opportunities and earning potential, which in turn affects the household's income and thus poverty status.

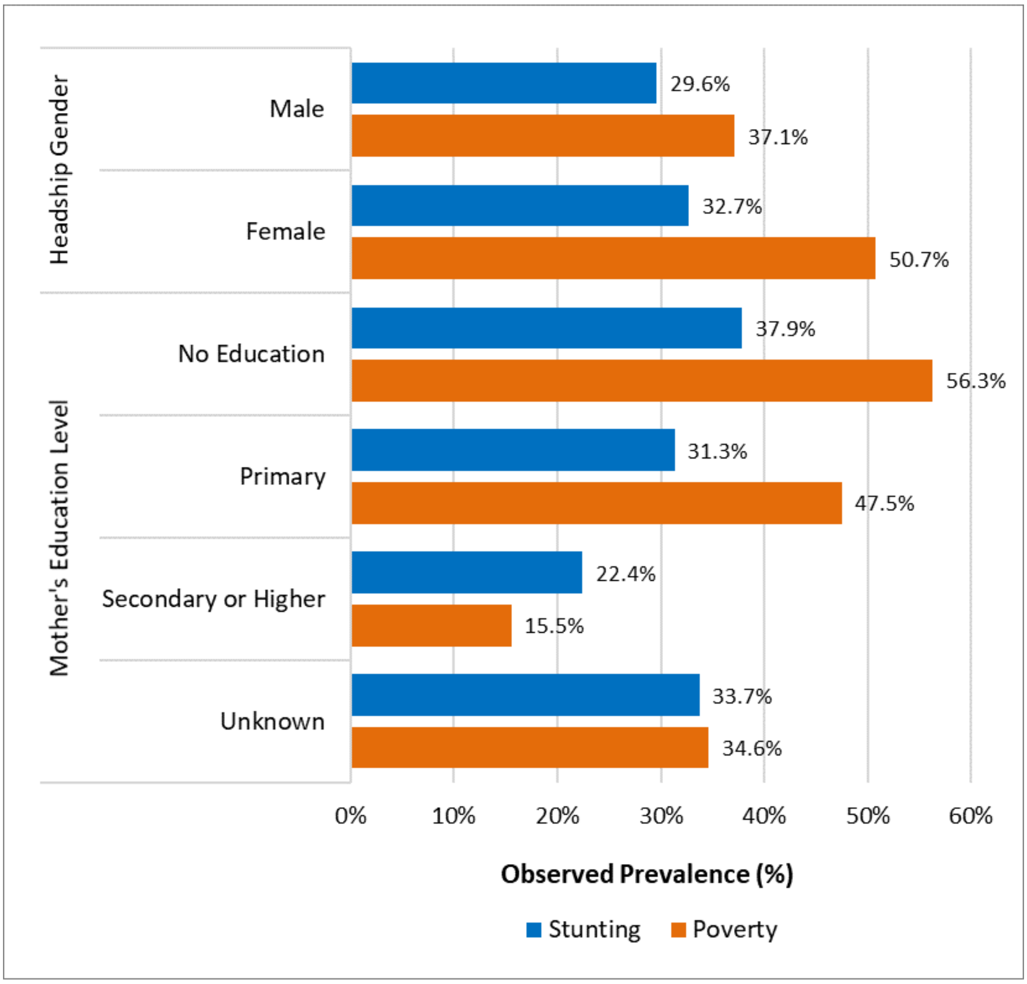


Figure 2.13: Observed prevalence of child poverty and stunting according to headship gender and mother's highest education level

2.5 Statistical Methods Appropriate for this Study

Both response variables of interest are binary. Thus, the logistic regression model, which is a class of the generalized linear model (GLM), would naturally be the selected statistical model for analysis. However, the data used in this study was obtained using a complex survey design, where the clusters selected to be included in the survey represent only a random sample of clusters. In addition, multi-stage sampling was utilised. Therefore, the method of analysis should account for this survey design, of which there are many (Heeringa et al., 2010). This study will follow a model-based approach which accounts for the variability in the responses that is attributable to each of the multiple levels of sampling. This variability exists due to the fact that children within the same households and clusters tend to be more alike than those from different households and clusters. Such a model-based approach includes the generalized linear mixed model (GLMM), which will be used in this study to fit the separate models for child poverty and malnutrition in children as well as the joint model between the two responses.

The following two chapters give an overview of the GLM and GLMM, respectively. In addition, Chapter 4 introduces the Bayesian method of parameter estimation, which is the method used in this study to fit the models.

Chapter 3

Generalized Linear Models

Generalized linear models (GLMs) extend the idea of conventional regression analysis, first introduced by Nelder & Wedderburn (1972). Models for the analysis of non-normal responses using non-linear models have a long history. These models are suitable when used along with exploratory variables that are categorical and continuous. The GLM includes a large class of statistical models used for relating non-normal responses to a linear combination of predictor variables, which can also include interaction terms (McCullagh, 2019). The parameters of the model provide a way to assess which exploratory variables are related to the response variables.

3.1 The Exponential Family

The exponential family consists of a set of distributions for both continuous and discrete random variables. The exponential family of distribution is a unifying concept, which brings together the commonly used probability distributions under a generalized framework (Wu, 2005). These include distributions such as: Normal, Binomial, Bernoulli, Poisson, Gamma, Multinomial and inverse Gaussian. A well known exception is Weibull distribution, which is a common distribution for survival (event history) data. If Y_i, $(i = 1, 2, ..., n)$ is a response variable from a distribution that is a member of the exponential family, then the probability density function for Y_i, $f(y_i; \theta_i, \phi)$, can be written as:

$$f(y_i; \theta_i, \phi) = \exp\left[\frac{y_i\theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right] \tag{3.1}$$

where θ_i is known as the canonical or natural parameter, $a_i(\phi) = \phi/w_i$ with w_i as a known weight depending on whether or not the data is grouped and ϕ is the dispersion parameter, $b(\cdot)$ and $c(\cdot)$ are some known functions. Jørgensen (1987) showed that if a response variable Y_i has a distribution belonging to the exponential family, then its mean and variance are

$$E(Y_i) = \mu_i = b'(\theta i) \tag{3.2}$$

$$\begin{aligned} Var(Y_i) &= a_i(\phi)b''(\theta_i) && (3.3) \\ &= \frac{\phi}{w_i}b''(\theta_i) && (3.4) \end{aligned}$$

where $b'(\theta i)$ and $b''(\theta_i)$ are the first and second derivatives of $b(\theta_i)$ with respect to θ_i. $b''(\theta_i)$ is a function of the mean, and is referred to as the variance function, denoted by $v(\mu_i)$. ϕ is constant over the observations, however w_i varies from observation to observation, hence a property of the GLM is a non-constant variance, which may vary across the observations (Nelder & Wedderburn, 1972). If $a_i(\phi) > 1$, the model is said to be overdispersed as $Var(Y_i) > v(\mu_i)$. Similarly, the model will be underdispersed if $a_i(\phi) < 1$.

Agresti & Kateri (2011) described three primary components of a GLM:

The Random Component
A GLM consists of a response variable, Y_i from the exponential family of distributions with n independent observations given by $y_1, y_2, , ..., y_n$.

The Systematic Component
This component relates a vector $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_n)'$ to a set of explanatory variables through a link function. Let $\boldsymbol{x}_i = (1, x_{1i}, ..., x_{pi})'$ be a $(p+1)$-dimensional vector of covariates and $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_p)'$ be a vector of the regression coefficients. The distribution of Y_i depends on $\boldsymbol{x}_i$ through the linear predictor, η_i, such that:

$$\begin{aligned} \eta_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi} \\ &= \boldsymbol{x}_i'\boldsymbol{\beta} \end{aligned}$$

The Link Function
The link function, given by $g(\mu_i)$, is a monotonic and differentiable function that describes how the mean, $E(Y_i) = \mu_i$, depends on the linear predictor. Thus, the GLM is generally defined as:

$$g(\mu_i) = \eta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi}$$

The inverse of the link function, $g^{-1}(\eta_i) = \mu_i$ is known as the mean function. Special types of link functions are those obtained directly from the natural parameter θ

of the exponential family. These link functions are regarded as canonical link functions such as the logit link for binary data, log link for count data and identity link for normal data. Table 3.1 shows some of the common link functions based on the distribution of the response variable.

Table 3.1: Some common distributions of exponential dispersion family with their link functions

Distribution	Link name	Inverse link	Variance function
Normal	Identity	η	ϕ
Binomial	Logit	$e^{\eta}/(1+e^{\eta})$	$\mu(1-\mu)/N$
Poisson	Logarithms	e^{η}	μ
Gamma	Reciprocals	$e^{-\eta}$	$\phi\mu^2$

3.2 Parameter Estimation

To fit a particular model, we need to find its parameter estimates and to perform this task for the GLM, the maximum likelihood (ML) procedure is commonly used.

The log-likelihood function for a single observation, y_i, is defined as

$$\ell_i = ln\, f(y_i; \theta_i, \phi) = \frac{y_i\theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi) \tag{3.5}$$

Since we assume that the observations $y_i, i = 1, \ldots, n$, are independent, the joint log-likelihood function is given by

$$\ell(\boldsymbol{\beta}, \boldsymbol{y}) = \sum_{i=1}^{n} \ell_i \tag{3.6}$$

The maximum likelihood estimate (MLE) of $\beta_j, j = 0, \ldots, p$, is therefore the solution to the following score equation

$$s(\beta_j) = \frac{\partial \ell_i}{\partial \beta_j} = \sum_{i=1}^{n} \frac{y_i - \mu_i}{a_i(\phi)v(\mu_i)} \times \frac{x_{ij}}{g'(\mu_i)} = 0 \tag{3.7}$$

where $g'(\mu_i) = \dfrac{\partial \eta_i}{\partial \mu_i}$ depends on the link function. The above equation is commonly represented as

$$s(\boldsymbol{\beta}) = \sum_{i=1}^{n} (y_i - \mu_i) W_i \frac{\partial \eta_i}{\partial \mu_i} x_{ij} = 0 \tag{3.8}$$

where W_i is is the iterative weights given by

$$W_i = \frac{1}{a_i(\phi)v(\mu_i)}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 \tag{3.9}$$

$$= \frac{1}{Var(Y_i)}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 \tag{3.10}$$

Therefore, solving Equation 3.8 above will give the MLE of $\boldsymbol{\beta}$. However, this score equation is a non-linear function of $\boldsymbol{\beta}$, therefore iterative procedures are required to solve them, such as Newton Raphson and Fisher Score (McCullagh & Nelder, 1989).

3.3 Logistic Regression

In the case of a binary response, the outcome can be coded as follows

$$Y_i = \begin{cases} 1 & \text{if an event is observed, e.g. a child is stunted or living under poverty} \\ 0 & \text{if an event is not observed, e.g. a child is not stunted or living under poverty} \end{cases}$$

Therefore, Y_i follows a Bernoulli distribution with $P(Y_i = 1) = \pi_i$ and $P(Y_i = 0) = 1 - \pi_i$. Thus, the mean and variance are given by

$$E(Y_i) = \pi_i \quad \text{and} \tag{3.11}$$

$$Var(Y_i) = \pi_i(1 - \pi_i) \tag{3.12}$$

The logistic regression model is given by

$$\text{logit}(\pi_i) = \ln\left(\frac{\pi_i}{1-\pi_i}\right) = \boldsymbol{x}_i'\boldsymbol{\beta} \tag{3.13}$$

where the left side of the equation is referred to as the logit link (Kutner et al., 2005). Therefore, the mean function of the logistic regression model is given by

$$E(Y_i) = \pi_i = \frac{\exp(\boldsymbol{x}_i'\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_i'\boldsymbol{\beta})} \tag{3.14}$$

The value of the link η_i is able to range freely while that of $E(Y_i) = \pi_i = \mu_i$ is restricted between 0 and 1. The MLE of $\boldsymbol{\beta}$ can once again be found using the iterative procedures, where the score $s(\boldsymbol{\beta})$ can be found as follows.

The probability distribution for a binary response is given by

$$f(y_i) = \pi_i^{y_i}(1-\pi_i)^{1-y_i} \tag{3.15}$$

This can be expressed in the following form

$$f(y_i) = \exp\left[y_i \ln\left(\frac{\pi_i}{1-\pi_i}\right) + \ln(1-\pi_i)\right] \tag{3.16}$$

The equation above is in the same form as Equation 3.1 with $a_i(\phi) = 1$, the dispersion parameter $\phi = 1$, $c(y_i, \phi) = 0$ and the canonical parameter $\theta_i = \ln\left(\frac{\pi_i}{1-\pi_i}\right)$. This gives $\pi_i = \frac{e^{\theta_i}}{1+e^{\theta_i}}$. Therefore, it follows $b(\theta_i) = \ln(1 + e^{\theta_i})$.

Since, for the logistic regression model, $E(Y_i) = \pi_i = \mu_i$, $Var(Y_i) = \mu_i(1-\mu_i)$, the link function can be represented as follows

$$\begin{aligned}\eta_i &= \ln\left(\frac{\mu_i}{1-\mu_i}\right) \\ &= \ln(\mu_i) - \ln(1-\mu_i)\end{aligned}$$

Thus, it follows

$$\begin{aligned}\frac{\partial\eta_i}{\partial\mu_i} &= \frac{\partial}{\partial\mu_i}\left[ln(\mu_i) - ln(1-\mu_i)\right] \\ &= \frac{1}{\mu_i} + \frac{1}{1-\mu_i} \\ &= \frac{1}{\mu_i\,(1-\mu_i)}\end{aligned}$$

Therefore,

$$\begin{aligned}W_i &= \frac{1}{Var(Y_i)}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 && (3.17) \\ &= \frac{1}{\mu_i(1-\mu_i)}\left[\mu_i\,(1-\mu_i)\right]^2 && (3.18) \\ &= \mu_i\,(1-\mu_i) && (3.19)\end{aligned}$$

Thus, the score $s(\boldsymbol{\beta})$ given in Equation 3.8 reduces to

$$s(\boldsymbol{\beta}) = \sum_{i=1}^{n}(y_i - \mu_i)\,x_{ij}$$

where μ_i, which is equal to π_i, is given by Equation 3.14.

A convenient property of the logistic regression is that the link function represents the log of the odds of an event of interest occurring, where $\frac{\pi_i}{1-\pi_i}$ is the odds of the event occurring. Therefore, e^{β_j} gives the odds ratio corresponding to a one unit increase in the corresponding explanatory variable x_{ij}, while all other explanatory variables remain unchanged. In general, the odds ratio for a k unit change in the explanatory variable is $e^{k\,\beta_j}$. This is a useful way in which to determine how much more likely an event of interest is to occur as one explanatory variable changes (Kutner et al., 2005).

While the logistic regression is a popular method to assess the relationship between a binary outcome and a set of explanatory variables, this method does not take into account the design of a study. In addition, the logistic regression model treats the effects of all the explanatory variables as fixed. However, due to the design of the study and the hierarchical nature of the data being used, it is necessary to include random effects in the model. Thus, the next chapter considers the generalized linear mixed model, which incorporates fixed as well as random effects.

Chapter 4

Generalized Linear Mixed Models

A generalized linear mixed model (GLMM) is one that can accommodate correlated data (Nelder & Wedderburn, 1972). The GLMM is an extension of the GLM that allows researchers to model non-normal and non-linear data which includes both random effects and fixed effects. We note that the MDHS data being analysed in this study consists of clusters and households within the cluster, and therefore possible correlations may exist in the responses. It is in this respect that we consider the GLMM, as it allows for the addition of random effects which can account for the possible correlation in the observations.

4.1 Model Formulation

On specifying the model, we begin with the conditional distribution of Y_{ij}, which is the j^{th} response, $j = 1, \dots, n_i$, from the i^{th} cluster, $i = 1, \dots, n$, given a vector $\boldsymbol{u}$ of random effects. The assumption is that, given a vector $\boldsymbol{u}$, the responses Y_{ij} are conditionally independent with a density from the exponential family (McCulloch & Neuhaus, 2005). That is

$$Y_{ij}|\boldsymbol{u} \sim \text{indep.} f(y_{ij}|\boldsymbol{u})$$

where

$$f(y_{ij}|\boldsymbol{u}) = \exp\left[\frac{y_{ij}(\theta_{ij}) - b(\theta_{ij})}{a_{ij}(\phi)} + c(y_{ij}, \phi)\right] \tag{4.1}$$

Functions b, c, and $a_{ij}(\phi)$ are defined as those in the GLM in Equation 3.1. The conditional mean of Y_{ij} is given by $E(Y_{ij}|\boldsymbol{u}) = \mu_{ij}$ and is modelled through a linear predictor, η_{ij}, containing fixed regression parameters $\boldsymbol{\beta}$, as well as cluster-specific parameters $\boldsymbol{u}_i$. That is

$$\begin{aligned}\eta_{ij} = g(\mu_{ij}) &= g\left[E(y_{ij}|\boldsymbol{u}_i)\right] \\ &= \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{u}_i \end{aligned} \tag{4.2}$$

where $\boldsymbol{x}_{ij}$ and $\boldsymbol{z}_{ij}$ are $(p+1)$-dimensional and q-dimensional vectors of known covariates corresponding to the fixed and random effects, $\boldsymbol{\beta}$ and $\boldsymbol{u}_i$, respectively. It is also assumed that the random effects $\boldsymbol{u}_i \sim \boldsymbol{MVN}(\boldsymbol{0}, \boldsymbol{G})$, where $\boldsymbol{G}$ depends on unknown variance components.

To estimate the parameters and variance components of the GLMM, two approaches are used: a Bayesian approach and the maximum likelihood approach. While the maximum likelihood approach, and variants of it such as restricted maximum likelihood (REML), is the most commonly used method of estimation, for GLMMs with high dimensional random effects or GLMMs with many random effects, this approach becomes computationally intensive, particularly with non-Gaussian responses (McCulloch & Neuhaus, 2005). However, simulation-based methods such as Markov chain Monte Carlo, specifically Gibbs sampling, and Monte Carlo EM have proven useful under these settings (McCulloch, 1997). The latter forms part of the Bayesian approach, which will be focussed on in this thesis.

4.2 Bayesian Approach to Parameter Estimation

In the Bayesian approach to estimation of the GLMM, there is a need to specify the prior densities and thereafter the posterior distribution can be found.

4.2.1 Introduction to Bayesian Methods

Unlike in the frequentist approach, Bayesian methods require prior information to estimate the posterior distribution. So, the challenge with the Bayesian approach is on estimating the posterior distribution, which requires integration of high-dimensional functions. There are several approaches to remedy this problem and they involve simulating realisations from the joint posterior distribution. In this thesis, the focus will be on the Markov chain Monte Carlo (MCMC) methods of simulating data.

The MCMC approach uses the previous sample values to randomly generate the next sample value, generating a Markov chain. The Bayesian analysis technique starts with an initial probability distribution for parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)'$, which is given by $p(\boldsymbol{\theta})$, referred to as the prior. This prior represents an assumption or previous knowledge about the parameters before observing the data. With observed data $\boldsymbol{y} = (y_1, \ldots, y_n)$, we then select a statistical model with likelihood function $p(\boldsymbol{y}|\boldsymbol{\theta})$. Thereafter, we combine the prior beliefs about $\boldsymbol{\theta}$ with the data $\boldsymbol{y}$ to give us a posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{y})$. The posterior distribution thus reflects the researcher's updated beliefs, and may be summarised in the form of posterior statistics, such as the posterior mean, median or mode, and standard deviation, and through the creation of Bayesian confidence intervals, known as credible intervals (Lesaffre & Lawson, 2012). These intervals have a natural interpretation. The Bayesian inference is derived from simple probability theory based on Bayes' Theorem. If A and B are any two events, then the conditional probability of A given B is derived as

$$P(A|B) = \frac{P(A,B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

If we replace B by observations $\boldsymbol{y}$, A by a parameter set $\boldsymbol{\theta}$ and probabilities by densities, the above equation results in the following relation

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{p(\boldsymbol{\theta}, \boldsymbol{y})}{p(\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{y})} \tag{4.3}$$

where $p(\boldsymbol{y}|\boldsymbol{\theta})$ represents the likelihood of $\boldsymbol{y}$ under the model and $p(\boldsymbol{\theta})$ is the prior density, or the density of $\boldsymbol{\theta}$ before observing the data $\boldsymbol{y}$. Furthermore, the denominator, $p(\boldsymbol{y})$, in Equation 4.3 is the marginal probability density of $\boldsymbol{\theta}$ such that

$$p(\boldsymbol{y}) = \int p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}$$

In addition, $p(\boldsymbol{y})$ does not depend on $\boldsymbol{\theta}$ and is regarded as a normalizing constant in order to ensure that the posterior distribution integrates to one. Therefore, an equivalent to Equation 4.3 is given by

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{4.4}$$

Thus, the posterior distribution is proportional to the likelihood function times the prior. Note, the likelihood $p(\boldsymbol{y}|\boldsymbol{\theta})$ may also be represented as $L(\boldsymbol{\theta}|\boldsymbol{y})$.

4.2.2 Prior Distributions

For a Bayesian approach, all parameters are treated as random variables and are assigned appropriate prior probability distributions. Therefore, a single value of a parameter is simply one possible realisation of the possible values of the parameter, the probability of which is defined by the prior distribution (Lawson, 2011). A prior distribution is usually assigned to a parameter before seeing the data, therefore it can be considered as providing additional data for a problem, which can result in improved estimations or identification of parameters. Thus, prior distributions and the likelihood provide two sources of information about the parameter to be estimated. The likelihood informs us about the parameter via the data, while the prior distributions inform via prior beliefs or assumptions (Lawson, 2011). For large amounts of data, the likelihood will contribute more to the posterior. However, when the data is lacking, the prior distributions will dominate the analysis. Different types of prior distributions can be specified depending on the prior knowledge available. Some of these are discussed below.

Improper Priors

It is possible that a specified prior distribution can be improper. Impropriety is a condition in which the integration of the prior distribution over the range of the parameter $\boldsymbol{\theta}$ is not finite. Thus, a prior distribution is considered improper if its normalizing constant is infinite. Although impropriety is a limitation of any prior distribution, it does not necessarily lead to impropriety of the posterior distribution. Therefore, the posterior distribution can be proper even with an improper prior specification (Lawson, 2013).

Non-informative Priors

A type of prior that is assumed to not make a strong preference over the observed values is regarded as non-informative prior. Thus, these priors are also known as vague or flat prior distributions. Intuitively, this means that choosing a non-informative prior distribution for the parameters in any posterior distribution will not have significant impact, compared to the likelihood of the data. The choice of non-informative prior can be made with some general understanding of the range and behavior of the parameters (Lawson, 2013). Also, we take note that variance parameters must have prior distributions on the positive real line. Plausible prior distributions in this case are often obtained in the gamma, inverse gamma, or uniform families.

Conjugate Prior

When choosing a prior distribution, a particular combination of the prior distribution and likelihood function can result to the same distribution family in the posterior $p(\boldsymbol{\theta}|\boldsymbol{y})$ as for the prior distribution $p(\boldsymbol{\theta})$. We call these prior and posterior distributions conjugate distributions, and the prior is called a conjugate prior for that particular likelihood. For example, the binomial likelihood with parameter $\boldsymbol{\theta}$ and the Beta prior distribution for $\boldsymbol{\theta}$ are conjugate. Thus, in this regard, the posterior distribution of $\boldsymbol{\theta}$ is also Beta distributed. The same holds for the Poisson likelihood, and the gamma prior distribution, as well as the normal likelihood with a normal prior distribution for the mean. Moreover, all members of the exponential family have conjugate priors (Lawson, 2011).

By examining the kernel of the prior likelihood product, the conjugacy prior can be identified. Thus, the prior likelihood product should be in a similar form to the prior distribution. In this regard, conjugacy always guarantees a proper posterior distribution.

4.2.3 Posterior Summary Measures

After obtaining the posterior distribution of a parameter $\boldsymbol{\theta}$, it is of interest to obtain summary measures of the posterior using representative values of the location and variability (Lesaffre & Lawson, 2012). The three most commonly used measures of location of $p(\boldsymbol{\theta}|\boldsymbol{y})$ include (i) the posterior mode; (ii) the posterior mean; and (iii) the posterior median.

The *posterior mode* is the value of $\boldsymbol{\theta}$ which maximises $p(\boldsymbol{\theta}|\boldsymbol{y})$, thus it is defined as

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\boldsymbol{y}),$$

which has the following properties:

- Maximization is required to find $\widehat{\boldsymbol{\theta}}$, therefore only $p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ is involved.
- For a flat prior distribution where $p(\boldsymbol{\theta}) \propto c$, the posterior mode is equal to the MLE of $\boldsymbol{\theta}$ from a frequentist approach using only the likelihood function as $p(\boldsymbol{\theta}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\theta})$.

The *posterior mean* is given by

$$\overline{\boldsymbol{\theta}} = \int \boldsymbol{\theta}\, p(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta}, \tag{4.5}$$

with the following properties:

- It is the value of $\boldsymbol{\theta}$ that minimizes the squared loss, given by

$$\int (\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})^2\, p(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta}$$

This means that $\overline{\boldsymbol{\theta}}$ is closest to all $\boldsymbol{\theta}$, as measured by the quadratic loss function $L(\boldsymbol{\theta}, \boldsymbol{\phi}) = (\boldsymbol{\theta} - \boldsymbol{\phi})^2$ weighted by the posterior distribution.

- To obtain the posterior mean, integration is performed twice: i) to determine the posterior distribution and ii) in Equation 4.5 above.

The *posterior median*, $\overline{\boldsymbol{\theta}}_M$, is defined by

$$0.5 = \int_{\overline{\boldsymbol{\theta}}_M} p(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta}, \tag{4.6}$$

with the following properties:

- The posterior median is obtained if the quadratic loss function in Equation 4.5 is replaced by $a|\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}|$, where $a > 0$.
- One integration and solving of an integral equation is required in obtaining the posterior median.

The most common posterior measure of variability is the *posterior variance*, given by

$$\overline{\boldsymbol{\sigma}}^2 = \int (\boldsymbol{\theta} - \overline{\boldsymbol{\theta}})^2 \, p(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta} \tag{4.7}$$

The posterior standard deviation, $\overline{\boldsymbol{\sigma}}$, is a useful measure of posterior uncertainty about the parameter of interest. However, to calculate the posterior variance/standard deviation, three integrals need to be evaluated (Lesaffre & Lawson, 2012).

4.2.4 The Markov Chain Monte Carlo Methods

A convenient way of obtaining summary measures of the posterior distribution, rather than evaluating several integrals, is to make use of Markov Chain Monte Carlo (MCMC) methods. MCMC comprises of a class of algorithms in order to directly sample from the posterior distribution. The summary measures are then estimated from those samples (Lesaffre & Lawson, 2012).

In this algorithm, we consider $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(T)}$ to be a sample of size T from the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{y})$. A Markov chain is a stochastic process in which future states are independent of past stages, given the present states (Ntzoufras, 2011). Based on the sequence $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(T)}$, the process can be specified by

$$P(\boldsymbol{\theta}^{(t+1)}|\boldsymbol{\theta}^{(t)}, \ldots, \boldsymbol{\theta}^{(1)}) = P(\boldsymbol{\theta}^{(t+1)}|\boldsymbol{\theta}^{(t)}),$$

which is the distribution of θ at time $t+1$, given all the proceeding θ values (for $t, t-1, ..., 1$), depends only on the value $\boldsymbol{\theta}^{(t)}$ at the immediate past time t.

In the limit, as $t \to \infty$, the distribution of $\boldsymbol{\theta}^{(t)}$ converges to its equilibrium, which is independent of the initial value of the chain $\boldsymbol{\theta}^{(0)}$. Thus, this condition occurs whenever the Markov chain is irreducible, aperiodic and positive-recurrent.

MCMC techniques work by building the Markov chain in a set of steps as outlined below:

1. Start with an initial value $\boldsymbol{\theta}^{(0)}$
2. Sequentially generate $\boldsymbol{\theta}^{(t+1)}|\boldsymbol{\theta}^{(t)}$ values from $p(\boldsymbol{\theta}|\boldsymbol{y})$ until the equilibrium distribution is attained.
3. Monitor the convergence of the algorithm using the convergence diagnostics. If convergence diagnostics fails, we generate more samples.
4. Cut off the first B observations. Then B is called the burn-in period; meaning the first B iteration values are eliminated from the sample to avoid the influence of initial values.
5. $\{\boldsymbol{\theta}^{(B+1)}, \boldsymbol{\theta}^{(B+2)}, ..., \boldsymbol{\theta}^{(T)}\}$ is considered as the sample of the posterior analysis.
6. Plot the posterior distribution. Histogram and Kernel plots of the distribution are the most frequenly used plots.
7. Lastly, obtain summary results for the posterior distribution.

These steps mentioned above are necessary when assessing the convergence diagnostics. Convergence diagnostics are used for checking whether or not convergence has been attained. The two most popular MCMC methods that are used are the Metropolis-Hasting algorithm and the Gibbs sampler (Lawson, 2011), which are considered below.

The Metropolis-Hastings Algorithm

This is a Markov Chain Monte Carlo (MCMC) approach that was first introduced by Metropolis et al. (1953) in a statistical physics context and further generalized by Hastings (1970) based on statistical problems.

Often, the strain that arises with the Monte Carlo integration is that of obtaining samples from some complex probability distribution $p(\boldsymbol{x})$. Attempts to solve this problem are the roots of MCMC methods. These methods seek attempts by mathematical physicists to integrate very complex functions by random sampling and the resulting Metropolis-Hastings (M-H) algorithm. In this algorithm, suppose we have $\boldsymbol{\theta}$ as a vector-valued parameter and $p(\boldsymbol{\theta}|\boldsymbol{y})$ the target posterior distribution from which we wish to generate the sample of size T. In Bayesian inference, the M-H algorithm proceeds through the following iterative steps:

1. Start with initial values $\boldsymbol{\theta}^{(0)}$
2. Repeat the following steps for $t = 1, 2, ..., T$
   - Set $\boldsymbol{\theta} = \boldsymbol{\theta}^{(t-1)}$
   - Generate new candidate values $\boldsymbol{\theta}'$ from a proposal distribution $q(\boldsymbol{\theta}'|\boldsymbol{\theta})$.
   - Calculate
   $$\alpha = \min\left(1, \frac{p(\boldsymbol{\theta}'|y)q(\boldsymbol{\theta}|\boldsymbol{\theta}')}{p(\boldsymbol{\theta}|\boldsymbol{y})q(\boldsymbol{\theta}'|\boldsymbol{\theta})}\right)$$
   - Update $\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}'$ with probability α or $\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^{(t-1)}$ with probability $1-\alpha$.

An imperative trait of this algorithm is the fact that its convergence depends on the proposal distribution. Moreover, in practice, it is vital to choose the proposal distribution with care, as poor choices considerably delay convergence to the target distribution.

Gibbs Sampler (GS)

This technique was originally introduced by Geman & Geman (1984), which is a special case of the Metropolis-Hasting algorithm that uses a proposal density $q(\boldsymbol{\theta}'|\boldsymbol{\theta}^{(t)})$. However, the desired sample is generated from the full conditional posterior distribution $p(\boldsymbol{\theta}_{-j}|\boldsymbol{\theta}_j, \boldsymbol{y})$, which is the proposal distribution, where $\boldsymbol{\theta}_{-j} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{j-1}, \boldsymbol{\theta}_{j+1}, \ldots, \boldsymbol{\theta}_p)'$ and $j = 1, 2, ..., p$.

One of the favourable characteristics of the Gibbs sampler is that in each step, the random values must be generated from a uni-dimensional distribution, for which a variety of computational tools exists. Often, the Gibbs sampler moves to a new value and does not require specification of a proposal distribution (Ntzoufras, 2011).

Gibbs sampler algorithm steps:

1. Start with initial values $\boldsymbol{\theta}^{(0)}$

2. Repeat the following steps for $t = 1, 2, ..., T$

   - Set $\boldsymbol{\theta} = \boldsymbol{\theta}^{(t-1)}$
   - For $j = 1, 2, ..., p$, update θ_j from $\theta_j \sim p(\theta_j|\boldsymbol{\theta}_{-j}, \boldsymbol{y})$
   - Set $\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}$ and save it as generated set of values at the $(t+1)^{th}$ iteration of the algorithm

Thus, for a given particular state of the chain $\boldsymbol{\theta}^{(t)}$, new parameter values can be generated by

$\theta_1^{(t)}$ from $p(\theta_1|\theta_2^{(t-1)}, \theta_3^{(t-1)}, ..., \theta_p^{(t-1)}, \boldsymbol{y})$
$\theta_2^{(t)}$ from $p(\theta_2|\theta_1^{(t)}, \theta_3^{(t-1)}, ..., \theta_p^{(t-1)}, \boldsymbol{y})$
$\theta_3^{(t)}$ from $p(\theta_3|\theta_1^{(t)}, \theta_2^{(t)}, \theta_4^{(t-1)}, ..., \theta_p^{(t-1)}, \boldsymbol{y})$

$\vdots$
$\vdots$
$\vdots$

$\theta_j^{(t)}$ from $p(\theta_j|\theta_1^{(t)}, \theta_2^{(t)}, ..., \theta_{j-1}^{(t)}, \theta_{j+1}^{(t-1)}, ..., \theta_p^{(t-1)}, \boldsymbol{y})$
$\vdots$
$\vdots$
$\vdots$

$\theta_p^{(t)}$ from $p(\theta_j|\theta_1^{(t)}, \theta_2^{(t)}, ..., \theta_{p-1}^{(t)}, \boldsymbol{y})$.

The conditional distributions $p(\theta_j|\theta_1^{(t)}, \theta_2^{(t)}, ..., \theta_{j-1}^{(t)}, \theta_{j+1}^{(t-1)}, ..., \theta_p^{(t-1)}, \boldsymbol{y})$ are called full conditional distributions because θ_j is conditioned on all other parameters (Lawson, 2011).

4.2.5 Assessing and Improving Convergence in MCMC Technique

MCMC techniques make use of diagnostics criteria to assess whether the iterative simulations have reached an equilibrium distribution of the Markov Chain or not. These methods are crucial as they help to decide how many iterations to use and represent the posterior density to ensure that the Markov Chain converges. Thus, once the posterior distribution has been determined, checking convergence becomes easy. Below we look at several ways to assess the efficiency and ways of improving the convergence within MCMC.

Visual examination of the trace plot is one of the useful ways to assess convergence, which is a plot of the sample values versus the sample number. The figure below displays some of the trace plots that can result.

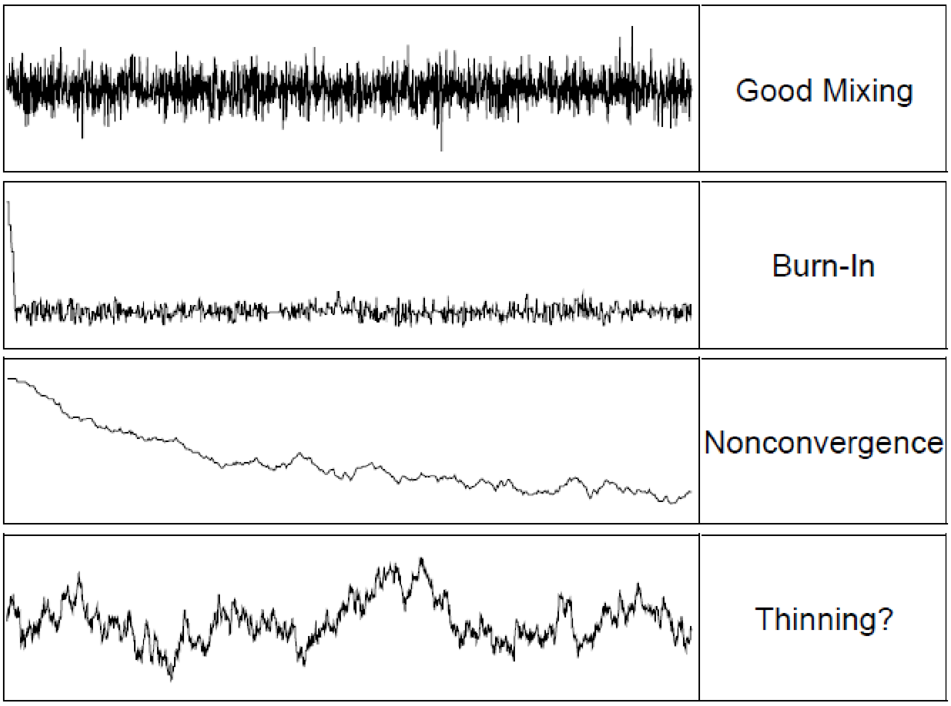


Figure 4.1: Types of trace plots

The efficiency of the MCMC algorithm can be assessed graphically; some of the methods are discussed below, as outlined by Lawson (2013) and Lesaffre & Lawson (2012):

- **Kernel Density Plots:** A more satisfactory density plot for a converged chain would look more like the bell-shaped density or indicative parameters whose marginal posterior densities are approximately normal
- **Autocorrelation Function Plot:** when future positions in the Markov chain are highly predictable from the current position, the posterior is slowly explored and one can be able to tell whether the chain has a mixing rate. A mixing rate is assessed by autocorrelations of different lags. If the mixing rate is low, then autocorrelation decreases slowly with increasing lag. The autocorrelation also indicates the minimum number of iterations needed for the Markov chain to "forget" its starting position.

- **Raftery and Lewis diagnostics:** The Raftery and Lewis diagnostic is used for detecting convergence to the stationarity distribution and evaluating accuracy of the estimated percentiles by the number of samples to reach the desired accuracy of the percentiles. This diagnostic is designed to test the number of iterations and burn-in needed by first running and taking a shorter pilot chain

- **Brooks-Gelman-Rubin diagnostic:** Brooks-Gelman-Rubin diagnostic is based on assessing the convergence of multiple chains through comparison of summary measures across chains. This diagnostic compares the within and between chain variances for each variable. When the chains have "mixed" (converged) the variance within each sequence and the variance between sequences for each variable will be approximately equal. However, this means that the ratio of the two variances converge to 1 if all chains being sampled are identically distributed. For poorly identified models, the variability of sampled parameter values between chains significantly exceeds the variability within one chain. The convergence of the Markov chain can be improved by standardising covariates and the use of unstructured random effects. Although, these methods are the most popular in assessing and improving convergence, we could also assess convergence by examining a time series plot or trace plots of the parameters. There is no guarantee that sampling using an MCMC algorithm will ensure or result to convergence to the posterior distribution.

4.2.6 Model Diagnostics and Selection

The most popular method used for model selection in Bayesian statistics is the deviance information criteria (DIC) that was proposed by Spiegelhalter et al. (2014). The DIC gives a measure that indicates both model fit and complexity and is defined as two times the mean of the deviance minus the deviance of the mean. It is based on the deviance and effective number of parameters in the model. The deviance of the posterior expected parameter, $\boldsymbol{\theta}$, can be defined as $D(\boldsymbol{\theta}) = -2 \ln p(\boldsymbol{y}|\boldsymbol{\theta})$ and is called the Bayesian deviance. Spiegelhalter et al. (2014) defined the effective number of parameters p_D in the model, which measure the model complexity which can be written as

$$p_D = \overline{D(\boldsymbol{\theta})} - D(\overline{\boldsymbol{\theta}}) \tag{4.8}$$

where $\overline{D(\boldsymbol{\theta})}$ is the posterior mean of $D(\boldsymbol{\theta})$ that measures the goodness of the model fit, while $D(\overline{\boldsymbol{\theta}})$ is the same as $D(\boldsymbol{\theta})$ evaluated at the posterior mean of the parameters. Thus, the DIC is defined as

$$DIC = D(\overline{\boldsymbol{\theta}}) + 2p_D \tag{4.9}$$

or equivalently,

$$DIC = \overline{D(\boldsymbol{\theta})} + p_D \tag{4.10}$$

One aspect of both DIC and p_D is that they can be easily calculated from the samples generated by MCMC simulation (Lesaffre & Lawson, 2012). Thus, smaller values of $\overline{D(\boldsymbol{\theta})}$ implies a good fit, whereas smaller values of p_D indicate model parsimony.

4.3 Joint Modelling using the GLMM Approach

Joint modelling of multiple outcomes allows for the correlation between the outcomes to be investigated and accounted for. The joint model optimally uses the available information from the multiple responses occurring simultaneously. In addition, the joint model has better control of type I error rates in multiple tests, thus it has possible gains in the efficiency of parameter estimates as opposed to separate models (Warton et al., 2015). There are various ways in which a joint model can be fitted, this includes the Plackett-Dale approach, GLMM approach, and probit normal approach, among others. In this study, we consider the GLMM approach to fitting a joint model.

For the purpose of this study, we will only consider the bivariate case with two binary responses as follows

$$Y_{i1} = \text{the poverty status of the } i^{th} \text{ child}$$
$$Y_{i2} = \text{the stunting status of the } i^{th} \text{ child}$$

The observed outcomes of these responses arise from a bivariate Bernoulli distribution, with π_{i1} equal to the probability that the i^{th} child is living under poverty, and π_{i2} equal to the probability that the i^{th} child is stunted. Therefore, for the bivariate response $\boldsymbol{Y}_i = (Y_{i1}, Y_{i2})'$, a GLMM of the following form is assumed

$$\boldsymbol{Y}_i = \boldsymbol{h}(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{u}_i) + \boldsymbol{\varepsilon}_i \tag{4.11}$$

where $\boldsymbol{h}(.) = \boldsymbol{g}^{-1}(.)$ is the inverse link function, the components of which depend on the link functions of the two responses, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)'$ contains the vectors of the fixed effects for each response, $\boldsymbol{u}_i = (\boldsymbol{u}_{i1}, \boldsymbol{u}_{i2})'$ contains the vectors of the random effects for each response, $\boldsymbol{X}_i = (\boldsymbol{X}_{i1}, \boldsymbol{X}_{i2})'$ and $\boldsymbol{Z}_i = (\boldsymbol{Z}_{i1}, \boldsymbol{Z}_{i2})'$ are the design matrices for the fixed and random effects, respectively, and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2})'$ is the residual error terms. Therefore, the model in Equation 4.11 is written in its most general form, as a decomposition in terms of the mean and error term, where both the mean and error term are allowed to change with the nature of the responses (Faes et al., 2008). It should be noted that the link function for each response is allowed to differ. In the case of this study, however, both responses use the logit link function.

Using this GLMM approach, the correlation between the outcomes can be modelled in two ways: (i) via a shared random effect, or (ii) via the residual variance of $\boldsymbol{Y}_i$:

(i) Shared Random Effect

In the case of accounting for the correlation between the response via a shared random effect, it follows

$$\boldsymbol{u}_i = \begin{pmatrix} \boldsymbol{u}_{i1} \\ \boldsymbol{u}_{i2} \end{pmatrix} \sim \text{i.i.d. } \mathbf{MVN}\left(\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma}_i = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right) \tag{4.12}$$

where $\boldsymbol{\Sigma}_i$ is the variance-covariance matrix of the joint model, Σ_{11} and Σ_{22} are the variance components for each response, and $\Sigma_{12} = \Sigma_{21}$ are the correlation components between the two responses. In this study, these correlation components represent the correlation between a child living under poverty and being stunted. If $\Sigma_{12} = \Sigma_{21} = 0$, then Equation 4.11 becomes equivalent to fitting two separate models for each response. Conditioned on the shared random effect, the joint responses of each child are assumed independent.

(ii) Residual Variance

If the correlation between the outcomes is modelled using the residual variance of $\boldsymbol{Y}_i$, it follows

$$\boldsymbol{\varepsilon}_i = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix} \sim \text{i.i.d. } \mathbf{MVN}\left(\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Omega}_i\right) \tag{4.13}$$

where the components of the variance-covariance matrix $\boldsymbol{\Omega}_i$ depend on the mean-variance relationship of the various responses and can contain, in addition, a correlation matrix $\boldsymbol{R}_i$ and an overdispersion parameter (Faes et al., 2008).

An appropriate covariance structure is selected for either $\boldsymbol{\Sigma}_i$ or $\boldsymbol{\Omega}_i$. Model fit statistics can be used to select the most appropriate structure for the model.

Covariance Structures

In the table below, we present some of the covariance structures that can be used for the structure of $\mathbf{\Sigma}_i$ or $\mathbf{\Omega}_i$.

Table 4.1: Some of the simpler covariance structures.

Structure	Description	Number of Parameters	$(i,j)^{th}$ element
AR(1)	Autoregressive lag 1	2	$\sigma_{ij} = \sigma^2\rho^{\lvert i-j\rvert}$
CS	Compound Symmetry	2	$\sigma_{ij} = \sigma_1 + \sigma^2 1(i=j)$
UN	Unstructured	$t(t+1)/2$	$\sigma_{ij} = \sigma_{ij}$
TOEP	Toeplitz	t	$\sigma_{ij} = \sigma_{\lvert i-j\rvert} + 1$
VC	Variance Component	q	$\sigma_{ij} = \sigma_k^2 1(i=j)$

First-order autoregressive
The autoregressive covariance structure has homogeneous variances and the correlation declines exponentially with distance (Kincaid, 2005; Littell et al., 2000). Hence, the two measurements that are right next to each other in time are considered to be correlated. Moreover, measurements are said to be less correlated as they get further apart.

$$AR(1) = \begin{pmatrix} 1 & \rho^1 & \rho^2 \\ \rho^1 & 1 & \rho^1 \\ \rho^2 & \rho^1 & 1 \end{pmatrix}$$

Compound Symmetry:
The exchangeable or compound symmetry correlation structure assumes a constant correlation between any two measurements within a subject, regardless of the time interval between the measurements. This structure may be more appropriate with data sets that have clustered observations such that there may be no logical ordering of observations within a cluster (Kincaid, 2005; Littell et al., 2000).

$$CS = \sigma^2 \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}$$

Unstructured
There are $t(t+1)/2$ parameters to be estimated when the correlation matrix is completely unspecified. The generality of this structure has drawback for having a large number of parameter to be tested.

$$UN = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}$$

The above structure can also be parameterised in terms of variances and correlations, with the covariance between two measurements given by $\sigma_i \sigma_j \rho_{ij}$, where the correlation $\rho_{ij} = 1$ when $i = j$ (Kincaid, 2005).

Toeplitz
The Toeplitz or the banded structure, its covariance depends only on the lag and not as a mathematical function with smaller number of parameters (Kincaid, 2005). The Toeplitz structure is similar to first-order autoregressive in that all measurement next to each other have the same correlation measurements which are two apart have same correlation different from the first. However, the correlations do not necessarily have the same pattern.

$$TOEP = \begin{pmatrix} \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_2 & \sigma_1 & \sigma^2 \end{pmatrix}$$

Variance Component
This structure is used when independence is assumed between the measurements. Thus, it is the simplest structure.

$$VC = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}$$

Finding the variance-covariance matrix

As interest is on the correlation structure between the two responses in the joint model, a general first-order approximate expression for the variance-covariance matrix of $\boldsymbol{Y}_i$ is derived as

$$\boldsymbol{V}_i = Var(\boldsymbol{Y}_i) \simeq \boldsymbol{\Delta}_i \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i' \boldsymbol{\Delta}_i' + \boldsymbol{\Omega}_i \tag{4.14}$$

where

$$\boldsymbol{\Delta}_i = \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\eta}_i}\right)\Big|_{\boldsymbol{u}_i=\mathbf{0}} \tag{4.15}$$

and

$$\boldsymbol{\Omega}_i \simeq \boldsymbol{\Phi}_i^{1/2}\boldsymbol{A}_i^{1/2}\boldsymbol{R}_i(\boldsymbol{\alpha})\boldsymbol{A}_i^{1/2}\boldsymbol{\Phi}_i^{1/2} \tag{4.16}$$

where $\boldsymbol{A}_i$ is a diagonal matrix containing the variance following the generalized linear specification of Y_{ik}, $k = 1, 2$, for a given random effects $\boldsymbol{u}_i = 0$. Further, $\boldsymbol{\Phi}_i$ is also a diagonal matrix but with the dispersion parameters along the diagonal. $\boldsymbol{\Omega}_i$ explains the variance-covariance in the residual error ε_i and $\boldsymbol{\Delta}_i\boldsymbol{Z}_i\boldsymbol{D}\boldsymbol{Z}_i'\boldsymbol{\Delta}_i'$ corresponds to the random effects structure in $h(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{u}_i)$ (Faes et al., 2008). When there are no random effects in Equation 4.11, a marginal model is obtained. When there are no residual correlations in $\boldsymbol{R}_i$, the resulting model is referred to as the conditional independence model or purely random-effects model (Faes et al., 2008).

The Bayesian methods described in Section 4.2 can be applied to estimate the parameters and variance components of the joint model.

In the next chapter, we now analyse child poverty and stunting separately and jointly using the methods discussed in this chapter.

Chapter 5

Application to the MDHS Data

In this section we consider Bayesian Multilevel models. The discussion begins with the seperate multilevel model followed by the joint multilevel generalized linear models. This approach is followed so that the distinction between the two types of models is clarified.

5.1 Results for the Separate Models

In this section, the results of the Bayesian GLMM fitted separately to child poverty and malnutrition in children are discussed. A Bayesian approach was adopted for estimation, and thus the *MCMCglmm* R package was used (Hadfield, 2010). Before the model was fitted, we checked for the presence of multicollinearity among the variables using the variance inflation factor (VIF). The VIF for each variable was between 1 and 3.5, suggesting that no significant multicollinearty was present between the independent variables.

To fit the model to each of the two responses, we utilised the normal distribution prior with a zero mean and a large variance (10^8) for the fixed effects parameters and an inverse-Wishart prior distribution for the variance components of the random effects. For a single variance component, the inverse Wishart takes two scalar parameters, (co)variance matrix V and degree of belief parameter nu. The distribution tends to a point mass on V as the degree of belief parameter, nu, goes to infinity. The distribution tends to be skewed to the right when nu is not very large, with a mode of $\frac{\mathrm{V{*}nu}}{\mathrm{nu}+2}$ and a mean of $\frac{\mathrm{V{*}nu}}{\mathrm{nu}-2}$ (Hadfield, 2010). When $\mathrm{V} = 1$, the resulting distributions are equivalent to inverse gamma distributions with shape and scale parameters set to $\frac{\mathrm{nu}}{2}$. In addition, as long as $\mathrm{V} > 1$ and $\mathrm{nu} > 1$, a proper prior is specified (Hadfield, 2010).

Three models were run for each response: i) a fixed effect model with no random effects, ii) a mixed effects model with only a cluster level random effect, iii) a mixed effects model with both a cluster level as well as household level nested within the cluster random effect. The DIC was used to choose the best fitting model between the three. However, the best fitting model for both responses was the mixed effects model with the two random effects (cluster and household level). Therefore, the results discussed will be based on this model.

Each model was fitted using 60 000 iterations with a burn-in phase of 10 000 and a thinning rate of 25. Diagnostic plots, which can be found in Appendix B, were used to monitor the convergence of the chain for each parameter estimate. Each model was run with varying values of the scale parameter, nu, in the inverse-Wishart prior for the random effects, however, the parameter estimates and their significance were not sensitive to these changes. The final models were based on $V = 1$ and $nu = 1$. Table 5.1 displays the posterior odds ratio (OR) estimates and their corresponding 95% credible intervals (CrI) for the separate model fitted to each response using the Bayesian GLMM.

Based on the results, gender and birth order were not significant factors for stunting in children, where the odds of stunting for females was not significantly different to that of males [OR: 0.906, CrI: 0.760 - 1.067], and there was no significant difference in the odds of stunting as the birth order increased by one unit [OR: 0.999, CrI: 0.994 - 1.005]. However, age in months had a strong effect on stunting. Children older than 6 to 11 months had significantly higher odds of being stunted. Children aged 36-47 months were almost 5 times more likely to be stunted compared to those aged 6-11 months [OR: 4.839, 95% CrI: 3.287 - 6.955]. As one would expect, the odds of stunting decreased with an improvement in anaemia status, where children with moderate anaemia, mild anaemia or no anaemia had significantly lower odds of stunting compared to those with severe anaemia. The results suggested that children from the Northern region were more likely to be stunted compared to those from the Central region [OR: 1.378, 95% CrI: 1.058 - 1.796]. However, the odds of stunting in children from the Southern region were not significantly different from that in children from the Central region [OR: 1.301, 95% CrI: 0.997 - 1.688]. The results also indicated that the more educated the mother, the less likely the child was to be stunted. However, there was only a significant difference in the odds of stunting for children whose mother had a secondary or higher education compared to no education [OR: 0.694, 95% CrI: 0.486 - 0.979]. There was no significant difference in the odds of stunting based on the source of drinking water in the household, type of place of residence, headship age and gender, and household size. However, the odds of stunting significantly decreased with an increase in one unit of the household's wealth index Z-score [OR: 0.666, 95% CrI: 0.570 - 0.776].

Table 5.1: Estimated odds ratios (OR) and their 95% credible intervals (CrI) based on the Bayesian GLMM fitted to malnutrition and poverty separately

	Stunted		Poverty	
	OR	**95% CrI**	**OR**	**95% CrI**
	Individual Level Factors			
Child Sex				
Male	1.000	Ref	-	-
Female	0.906	(0.760 - 1.067)	-	-
Child Age (Months)				
$6-11$	1.000	Ref	-	-
$12-23$	4.151	(2.894 - 5.913)*	-	-
$24-35$	2.913	(1.987 - 4.177)*	-	-
$36-47$	4.839	(3.287 - 6.955)*	-	-
$48-59$	3.818	(2.657 - 5.636)*	-	-
Anaemia Level				
Severe	1.000	Ref	-	-
Moderate	0.513	(0.263 - 0.920)*	-	-
Mild	0.391	(0.204 - 0.718)*	-	-
Not Anemic	0.319	(0.178 - 0.623)*	-	-
Birth Order	0.999	(0.994 - 1.005)	-	-
	Household Level Factors			
Head of Household Sex				
Male	1.000	Ref	1.000	Ref
Female	1.113	(0.914 - 1.363)	2.331	(1.919 - 2.822)*
Region				
Central region	1.000	Ref	1.000	Ref
Northern region	1.378	(1.058 - 1.796)*	3.611	(2.579 - 5.002)*
Southern region	1.301	(0.997 - 1.688)	2.236	(1.606 - 3.045)*
Mother's Education Level				
No education	1.000	Ref	1.000	Ref
Primary	0.819	(0.615 - 1.072)	0.722	(0.560 - 0.943)*
Secondary or Higher	0.694	(0.486 - 0.979)*	0.139	(0.098 - 0.194)*
Unknown	0.980	(0.509 - 1.866)	0.547	(0.290 - 1.000)
Source of Water				
Unprotected water	1.000	Ref	1.000	Ref
Protected water	0.872	(0.667 - 1.103)	0.526	(0.408 - 0.680)*
Tap water	0.872	(0.531 - 1.093)	0.089	(0.063 - 0.128)*
Type of Residence				
Urban	1.000	Ref	1.000	Ref
Rural	1.017	(0.735 - 1.442)	8.726	(5.811 - 13.647)*
Headship Age	1.006	(0.998 - 1.015)	0.981	(0.973 - 0.989)*
Size of Household	1.000	(0.918 - 1.087)	0.809	(0.771 - 0.850)*
Household Wealth Index Score	0.666	(0.570 - 0.776)*	-	-

* significant at a 5% level of significance

\- indicates not applicable to the model

All the household level factors considered for the child poverty model were significantly associated with the response. Poverty was more likely to occur in households headed by females [OR: 2.331, 95% CrI: 1.919 - 2.822], households in rural areas [OR: 8.726, 95% CrI: 5.811 - 13.647] and households in the Northern region [OR: 3.611, 95% CrI: 2.579 - 5.002] or Southern region [OR: 2.236, 95% CrI: 1.606 - 3.045] compared to those in the Central region of Malawi. However, the odds of poverty decreased with an increase in the child's mother's education level [OR: 0.722, 95% CrI: 0.560 - 0.943 for primary and OR: 0.139, 95% CrI: 0.098 - 0.194 for secondary or higher], as well as an improvement of source of drinking water [OR: 0.526, 95% CrI: 0.408 - 0.680 for protected water and OR: 0.089, 95% CrI: 0.063 - 0.128 for tap water]. A one unit increase in the age of the head of household [OR: 0.981, 95% CrI: 0.973 - 0.989] and the number of members in the household [OR: 0.809, 95% CrI: 0.771 - 0.850] resulted in a significant decrease in the odds of poverty.

Based on Table 5.2, there was considerable variability between households for the poverty model with the variance component for this random effect estimated at 165.1. The estimates of both variance components for both models are relatively far from zero, therefore confirming the need to account for these sources of variability in the models.

Table 5.2: Covariance parameter estimates for the separate models with two random effects

Covariance Parameter	Stunting	Poverty
Cluster	0.2094	16.98
Household nested within cluster	1.396	165.1

5.2 Results for the Joint Model

In this section, we discuss the results of the joint GLMM fitted to child poverty and malnutrition in children using the MDHS data. This was achieved by once again making use of the *MCMCglmm* package in R, which uses the Gibbs sampler. The model was fitted using 60 000 iterations with a burn-in phase of 10 000 and a thinning rate of 25. Diagnostic plots were again used to monitor the convergence of the chain for each parameter estimate.

All the individual and household variables were incorporated into the joint model, excluding household wealth index, which the poverty outcome was based on. In addition, the two random effects included in the separate models from the previous section were also included in the joint model, that being the cluster and household level random effects. The overall intercept of the joint model was removed in order to be able to directly compare the parameter estimates of the joint model. Rather, an intercept for each response was incorporated. The correlation between the two responses on each child was accounted for via the residual variance of the model. Various covariance structures were fitted and the DIC of the models were compared. The unstructured covariance structure produced the lowest DIC and thus was selected. This covariance structure has the following form

$$\boldsymbol{\Omega} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$$

Once again, we utilised the normal distribution prior with a zero mean and a large variance (10^8) for the fixed effects parameters and an inverse-Wishart prior distribution for the variance components of the random effects and error, with V $= 1$ and nu $= 1$. The results from the joint model of poverty and malnutrition of children revealed a positive interdependence between the two outcome variables with an estimated correlation of 0.1 (Table 5.3). All the variance components appeared to be significant at 5% level of significance. The variability between household was again fairly high at 17.76.

Table 5.3: Covariance Parameter Estimates for the Joint Model

Covariance Parameter	Estimate
Var(Household nested within cluster)	17.76
Var(Cluster)	1.072
Var(Malnutrition)	3.702
Var(Poverty)	1.231
Correlation between poverty and malnutrition	0.10

Table 5.4 presents the odds ratio estimates and their 95% credible intervals (CrI) for the joint model. Based on the results, the joint model again indicated that age (in months) was one of the major predictors of stunting, where children that were aged 12-23 months [OR: 5.713, 95% CrI: 3.673 - 9.395], 24-35 months [OR: 3.702, 95% CrI: 2.382 - 5.837], 36-47 months [OR: 6.484, 95% CrI: 4.070 - 10.587] and 48-59 [OR: 4.896, 95% CrI: 3.122 - 7.916] had significantly higher odds of being stunted compared to children aged 6-11 months. A significant association between anaemia level and stunting was also suggested by the joint model. The model showed that children with moderate anaemia [OR: 0.434, 95% CrI: 0.205 - 0.910], mild [OR: 0.308, 95% CrI: 0.142 - 0.656] and no anaemia [OR: 0.242, 95% CrI: 0.113 - 0.525] had significantly lower odds of being stunted compared to children with severe anaemia. Higher odds of stunting in children were suggested for those residing in households headed by females compared to those residing in households headed by males [OR: 1.254, 95% CrI: 1.004 - 1.584]. The model also revealed significantly higher odds ratio of stunting for children that resided in the Northern region [OR: 1.650, 95% CrI: 1.143 - 2.357] and Southern region [OR: 1.485, 95% CrI: 1.048 - 2.122] compared to children that resided in the Central region. The mother's education level was also observed as a strong predictor of stunting, where the model suggested that the odds tend to decrease with an increase in mother's education level. Although, only children with mothers who had a secondary or higher education had significantly different odds of stunting compared to those with mothers who had no education [OR: 0.545, 95% CrI: 0.364 - 0.811]. Furthermore, children residing in rural areas had a considerably higher odds of stunting compared to those in urban areas [OR: 1.642, CrI: 1.098 - 2.479]. Children in household with tap water had a significantly lower odds of stunting compared to those in households with unprotected water sources [OR: 0.605, 95% CrI: 0.395 - 0.920]. There was no significant different in the odds of stunting for children in households with protected or unprotected water. No significant association was found between stunting and the child's gender, birth order, headship age and size of the household.

Moreover, the results from the joint model also confirmed that the factors child's age and child's sex had no significant association with poverty, which is not surprising as these child level factors are not likely to have any impact on their poverty status. Children residing in households headed by females were more than twice as likely to be living under poverty compared to those residing in households headed by males [OR: 2.329, 95% CrI: 1.925 - 2.839]. Further [OR: 2.296, 95% CrI: 1.659 - 3.174], the model revealed that while children from the Southern region were more likely to be poor than compared to those in the Central region, children from the Northern region were almost 4 times more likely to be striked with poverty. The odds of living under poverty were 46.1% lower for children residing in households that used protected water compared to those residing in households that used unprotected water [OR: 0.539, 95% CrI: 0.416 - 0.695]. The odds of living under poverty decreased even more for those living in households with tap water [OR: 0.185, 95% CrI: 0.127 - 0.269]. The anaemia status of a child was significantly associated with the odds of them living under poverty, where children that had moderate anaemia [OR: 0.446, 95% CrI: 0.237 - 0.844], mild anaemia [OR: 0.298, 95% CrI: 0.208 - 0.756] or no anaemia [OR: 0.298, 95% CrI: 0.155 - 0.564] had significantly lower odds of living under poverty compare to those that had severe anaemic level. The joint model indicated that children in the rural areas were almost 9 times more likely to be living under poverty compared to those in urban areas [OR: 8.987, 95% CrI: 5.924 - 13.814]. However, their odds decreased with an increase in the headship's age [OR: 0.980, 95% CrI: 0.972 - 0.988]. While the odds decreased with an increase in the child's birth order [OR: 0.999, 95% CrI: 0.994 - 1.005] and household size [OR: 0.813, 95% CrI: 0.936 - 1.045], these factors did not appear to be significantly associated with the odds of living under poverty.

Table 5.4: Estimated odds ratios (OR) and their 95% credible intervals (CrI) based on the Bayesian GLMM jointly fitted to malnutrition and poverty

	Stunted		Poverty	
	OR	**95% CrI**	**OR**	**95% CrI**
		Individual Level Factors		
Child Sex				
Male	1.000	Ref	1.000	Ref
Female	0.908	(0.745 - 1.106)	1.073	(0.910 - 1.277)
Child Age (Months)				
$6-11$	1.000	Ref	1.000	Ref
$12-23$	5.713	(3.673 - 9.395)*	1.065	(0.769 - 1.469)
$24-35$	3.702	(2.382 - 5.837)*	1.031	(0.735 - 1.424)
$36-47$	6.484	(4.070 - 10.587)*	1.095	(0.798 - 1.522)
$48-59$	4.896	(3.122 - 7.916)*	0.964	(0.692 - 1.337)
Anaemia Level				
Severe	1.000	Ref	1.000	Ref
Moderate	0.434	(0.205 - 0.910)*	0.446	(0.237 - 0.844)*
Mild	0.308	(0.142 - 0.656)*	0.401	(0.208 - 0.756)*
Not Anaemic	0.242	(0.113 - 0.525)*	0.298	(0.155 - 0.564)*
Birth Order	0.998	(0.991 - 1.005)	0.999	(0.994 - 1.005)
		Household Level Factors		
Head of Household Sex				
Male	1.000	Ref	1.000	Ref
Female	1.254	(1.004 - 1.584)*	2.329	(1.925 - 2.839)*
Region				
Central region	1.000	Ref	1.000	Ref
Northern region	1.650	(1.143 - 2.357)*	3.693	(2.637 - 5.187)*
Southern region	1.485	(1.048 - 2.122)*	2.296	(1.659 - 3.174)*
Mother Education Level				
No education	1.000	Ref	1.000	Ref
Primary	0.808	(0.588 - 1.117)	0.731	(0.558 - 0.947)*
Secondary or Higher	0.545	(0.364 - 0.811)*	0.134	(0.093 - 0.191)*
Don't know	0.995	(0.472 - 2.104)	0.579	(0.312 - 1.089)
Source of Water				
Unprotected water	1.000	Ref	1.000	Ref
Protected water	0.833	(0.613 - 1.124)	0.539	(0.416 - 0.695)*
Tap water	0.605	(0.395 - 0.920)*	0.185	(0.127 - 0.269)*
Other	0.235	(0.014 - 3.082)	0.328	(0.048 - 2.237)*
Type of Residence				
Urban	1.000	Ref	1.000	Ref
Rural	1.642	(1.098 - 2.479)*	8.987	(5.924 - 13.814)*
Headship Age	1.006	(0.997 - 1.015)	0.980	(0.972 - 0.988)*
Size of Household	0.988	(0.936 - 1.045)	0.813	(0.936 - 1.045)

* significant at a 5% level of significance

Chapter 6

Discussion and Conclusion

In this study, we made use of a generalized linear mixed model with cluster level and household random effects to fit separate as well as joint models to child poverty and malnutrition in children, as measured by stunting. This was done by making use of data from the 2016 Malawi Demographic and Health Survey, which consisted of a multistage cluster sampling. The study aimed to determine the significant socio-economic and demographic risk factors associated with child poverty and malnutrition in children aged 6 to 59 months in Malawi, separately and jointly. Furthermore, a Bayesian method of estimation was used due to the complexity of the models. The results of the separate models were obtained in a matter of minutes, and the joint model took approximately 45 minutes to run.

The main factors that were found to be significantly associated with poverty and malnutrition in children from the joint model were the child's age, anaemia status, mother's education level, headship's sex and age, region of Malawi, source of drinking water, and type of residence. In comparing these results to that of the separate models, only the household size was no longer significantly associated with poverty in the joint model. Furthermore, in the joint model, source of drinking water and type of place of residence were significantly associated with stunting in children, however these factors were insignificant in the separate model for stunting. This could be as a result of the household's wealth index Z-score, which was included in the separate model only, having more of a pronounced effect on stunting. The joint model produced fairly similar parameter estimates to the separate models, however with some wider confidence intervals.

This study revealed a positive correlation between poverty and malnutrition in children, however, the correlation was fairly weak. This suggests that there are other factors that contribute more considerably to each response. While the joint model does not indicate any direction of association between the two responses, it is useful in estimating and controlling for the degree of association and it thus has better control over type I error rates. The joint model is believed to have more accurate parameter estimates and the ability to answer intrinsically multivariate questions (Habyarimana et al., 2016).

Both the separate and joint models showed a decrease in the odds of poverty and malnutrition as the mother's education level increased. This is unsurprising as educated individuals are generally more aware of health related issues. Many studies, such as Pettersson & Enström (2016) and Vorster (2010), have strongly emphasised the importance of educating parents about the implications of malnutrition, especially mothers. This is crucial because when mothers are knowledgeable about malnutrition, it reduces the risk of their children developing malnutrition disorders/diseases, such as Kwashiorkor and Marasmus. Our findings revealed that source of drinking water was one of the significant factors associated with malnutrition and poverty jointly. Our findings reaffirmed the findings of Pronczuk & Surdu (2008), who claimed that more than one billion people lack access to improved water sources worldwide. Developing countries like Malawi, are more vulnerable to poor access to improved water sources and consequently, this may result in severe and irreversible life threatening diseases. The child's anaemia status was significantly associated with poverty and malnutrition jointly. This significant association suggests that children living under poverty are more than likely suffering from anaemia due to the poor resources available to them. This can also be as a result of affordability and access to quality, nutritional food sources.

In developing countries like Malawi, where over half of all children are malnourished, it is important that the government implements targeted and concerted actions in the areas of health, access to basic services, education, and specific nutritional interventions in order to reduce the predicament of malnutrition. Studies indicate that the following types of interventions are most likely to be successful in effort to lower the level of malnutrition among Malawi children and poverty of households:

Household-level programs: Implementation of nutritional educational programs at the household level can help with accelerating the decline of malnutrition of children in Malawi. They should emphasize the importance of a more balanced diet.

Infants extra support: Government and NGOs should donate to infants micronutrient supplements containing the right vitamins, as malnutrition does its greatest harm during the early stages of birth.

Improve education of women: Again, in order to fight the predicament of poverty and malnutrition in children, it should be worthwhile to uplift the education of women in Malawi. Women are proven to be the best care takers of children.

Some of the limitations of this study include that of the data being from a cross-sectional design, thus the casual relationship between the variables could not be determined. In addition, the data was limited with the variables available, such as mother's anaemia status, age, smoking status, and the head of household's education level, among others which have been considered in other studies.

Future directions of this study include incorporating additional anthropometric measurements as responses into the joint model. In addition, the spatial effect on poverty and malnutrition will be considered in order to account for spatial autocorrelation that may be present.

References

Addae-Korankye, A. (2014). Causes of poverty in Africa: a review of literature. *Am. Int. J. Soc. Sci*, *3*, 147–153.

Agresti, A., & Kateri, M. (2011). *Categorical data analysis*. Springer.

Ajakaiye, D., & Adeyeye, V. (2001). Concepts, measurement and causes of poverty. *Central Bank of Nigeria Economic and Financial Review*, *39*(4), 8–44.

Amcoff, S. (1981). The impact of malnutrition on learning situation.

Birdsong, K. (2016). 10 facts about how poverty impacts education. `https://www.scilearn.com/blog/ten-facts-about-how-poverty-impacts-education`. Online accessed: May 2018.

Blössner, M., De Onis, M., Prüss-Üstün, A., Campbell-Lendrum, D., Corvalán, C., & Woodward, A. (2005). Quantifying the health impact at national and local levels. *WHO, Geneva*.

Chinyoka, K. (2014). Impact of poor nutrition on the academic performance of grade seven learners: A case of Zimbabwe. *International journal of learning and development*, *4*(3), 73–84.

Croft, T. N., Marshal, A. M. J., Allen, C. K., et al. (2018). Guide to DHS Statistics. Rockville, Maryland, USA: ICF.

Cross-Border Road Transport Agency (2016). Malawi country profile report. Tech. rep.

Davids, Y. D., & Gouws, A. (2013). Monitoring perceptions of the causes of poverty in south africa. *Social Indicators Research*, *110*(3), 1201–1220.

Demissie, S., & Worku, A. (2013). Magnitude and factors associated with malnutrition in children 6-59 months of age in pastoral community of Dollo Ado district, Somali region, Ethiopia. *Sci J Public Health*, *1*(4), 175–83.

eschool Today (2008). Types of poverty. `http://www.eschooltoday.com/poverty-in-the-world/types-of-poverty.html`. Online accessed: May 2018.

Essay Professors (2006). But can it be argued that poverty is the root of crime essay paper online. `https://essaysprofessors.com/samples/argumentative/can-it-be-argued-that-poverty-is-the-root-of-crime.html`. Online accessed: June 2018.

Faes, C., Aerts, M., Molenberghs, G., Geys, H., Teuns, G., & Bijnens, L. (2008). A high-dimensional joint model for longitudinal outcomes of different nature. *Statistics in medicine*, *27*(22), 4408–4427.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6), 721–741.

Greever, S. (2014). Poverty in education. *Online Submission*.

Habyarimana, F., Zewotir, T., & Ramroop, S. (2016). Joint modeling of poverty of households and malnutrition of children under five years from Demographic and Health Survey: Case of Rwanda. *8*(2), 108–114.

Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software*, *33*(2), 1–22.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications.

Health Central (2016). Poverty reduces life expectancy about 10 years. `https://www.healthcentral.com/article/poverty-reduces-life-expectancy-about-10-years`. Online accessed: June 2018.

Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). *Applied Survey Data Analysis*. Statistics in the Social and Behavioral Sciences Series. Chapman & Hall/CRC. Taylor & Francis Group, LLC.

Hunt, J. M. (2005). The potential impact of reducing global malnutrition on poverty reduction and economic development. *Asia Pacific Journal of Clinical Nutrition*, *14*.

International Monetary Fund (2017). Malawi Economic Development Document.

Investopedia (2018). Poverty. `https://www.investopedia.com/terms/p/poverty.asp`. Online accessed: June 2018.

Jamison, D. T. (1986). Child malnutrition and school performance in China. *Journal of development economics*, *20*(2), 299–309.

Jørgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society: Series B (Methodological)*, *49*(2), 127–145.

Kincaid, C. (2005). Guidelines for selecting the covariance structure in mixed model analysis, paper 198-30 in proceedings of the thirtieth annual sas users group conference. *Inc., Cary, North Carolina*.

Kutner, M., Nachtsheim, C., & Neter, J. (2005). *Applied linear statistical models*. Statistics in the Social and Behavioral Sciences Series. McGraw-Hill/Irwin, 5 ed.

Lawson, A. B. (2011). *Bayesian Disease Mapping*. Chapman & Hall/CRC, 3 ed.

Lawson, A. B. (2013). *Bayesian disease mapping: hierarchical modeling in spatial epidemiology*. Chapman and Hall/CRC.

Lesaffre, E., & Lawson, A. B. (2012). *Bayesian biostatistics*. John Wiley & Sons.

Littell, R. C., Pendergast, J., & Natarajan, R. (2000). Modelling covariance structure in the analysis of repeated measures data. *Statistics in medicine*, *19*(13), 1793–1819.

Live Strong (2017). How does malnutrition affect health? `https://www.livestrong.com/article/18046-signs-symptoms-malnutrition/`. Online accessed: June 2018.

Marini, A., & Gragnolati, M. (2003). Malnutrition and Poverty in Guatemala. World Bank Policy Research Working Paper 2967.

McCullagh, P. (2019). *Generalized linear models*. Routledge.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*, vol. 37. CRC press.

McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association*, *92*(437), 162–170.

McCulloch, C. E., & Neuhaus, J. M. (2005). Generalized linear mixed models. *Encyclopedia of biostatistics*, *4*.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of Chemical Physics*, *21*(6), 1087–1092.

Mok, T. Y., Gan, C., & Sanyal, A. (2007). The determinants of urban household poverty in Malaysia. *Journal of Social Sciences*, *3*(4), 190–196.

Mukherjee, S., & Benson, T. (2003). The Determinants of Poverty in Malawi, 1998. *World Development*, *31*(2), 339–358.

Mzumara, B., Bwembya, P., Halwiindi, H., et al. (2018). Factors associated with stunting among children below five years of age in Zambia: evidence from the 2014 Zambia demographic and health survey. *BMC Nutrition*, *4*(51).

National Statistical Office (NSO) [Malawi] and ICF (2017). Malawi Demographic and Health Survey 2015-16. Zomba, Malawi, and Rockville, Maryland, USA. NSO and ICF.

Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, *135*(3), 370–384.

Ntzoufras, I. (2011). *Bayesian modeling using WinBUGS*, vol. 698. John Wiley & Sons.

Nursing Times (2009). Malnutrition. `https://www.nursingtimes.net/archive/malnutrition-20-05-2009/`. Online accessed: June 2018.

Omole, D. A. (2003). *Poverty in Lesotho: a case study and policy options*. Department of Economics, National University of Lesotho.

Oxford Handbooks Online (2017). Poverty and crime. `http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199914050.001.0001/oxfordhb-9780199914050-e-28`. Online accessed: May 2018.

Pettersson, C., & Enström, F. (2016). Prevention of malnutrition in south africa among children.

Phipps, S. (2003). Poverty and health, cphi collected papers.

Poverties (2011). Effects of poverty on society, health, children and violence. `https://www.poverties.org/blog/effects-of-poverty`. Online accessed: May 2018.

Project, T. B. (2014). Five devastating effects of poverty. `https://borgenproject.org/5-effects-poverty/`. Online accessed: May 2018.

Project Rise (2018). The two-way street between poverty and education. `https://projectrise.news24.com/two-way-street-poverty-education/`. Online accessed: May 2018.

Pronczuk, J., & Surdu, S. (2008). Children's environmental health in the twenty-first century. *Annals of the New York Academy of Sciences*, *1140*(1), 143–154.

Rajkumar, A. S., Gaukler, C., & Tilahun, J. (2011). *Combating Malnutrition in Ethiopia: An evidence-based approach for sustained results*. World Bank Publications.

Richmond Vale Academy (2016). Effects of poverty. `http://richmondvale.org/effects-of-poverty/`. Online accessed: May 2018.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *76*(3), 485–493.

The World Bank (2014). Poverty and health.

The World Bank (2018). Poverty. `https://www.worldbank.org/en/topic/poverty/overview`. [Online accessed: June 2018].

UK Essays (2015). The effect of poverty on health. `https://www.ukessays.com/essays/sociology/the-effect-of-poverty-on-health.php`. Online accessed: June 2018.

Van der Berg, S. (2008). Poverty and education. *Education policy series*, *10*, 28.

Vorster, H. H. (2010). The link between poverty and malnutrition: A South African perspective. *Health SA Gesondheid*, *15*(1).

Warton, D. I., Blanchet, F. G., OHara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F. K. (2015). So many variables: joint modeling in community ecology. *Trends in Ecology & Evolution*, *30*(12), 766–779.

Webb, P., & Bhatia, R. (2005). Manual: Measuring and interpreting malnutrition and mortality. *Nutr Serv WFP Rome*.

Weiss, T. C. (2016). Health consequences of malnutrition. Tech. rep.

Wood, A., & Mayer, J. (2006). Malawi poverty and vulnerability assessment: Investing in our future. *Washington DC: World Bank*.

World Health Organization (2005). Dac guidelines and reference series poverty and health.

World Health Organization (2006). *WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development*.

World Health Organization (2017). Global Nutrition Monitoring Framework: operational guidance for tracking progress in meeting targets for 2025. Geneva.Licence: CC BY-NC-SA 3.0 IGO.

World Health Organization (2018a). Child malnutrition. `https://www.who.int/gho/child-malnutrition/en/`. Online accessed May 2018.

World Health Organization (2018b). Maternal, newborn, child and adolescent health. `https://www.who.int/maternal_child_adolescent/topics/child/malnutrition/en/`. Online accessed: May 2018.

Wu, Z. (2005). Generalized linear models in family studies. *Journal of Marriage and Family*, *67*(4), 1029–1047.

Wu House, G. Z. (1992). Rapid population growth and its implication for Malawi. *malawi medical journal*.

Yimer, G. (2000). Malnutrition among children in Southern Ethiopia: Levels and risk factors. *Ethiopian Journal of Health Development*, *14*(3).

Appendix A

In this section, we present a set of codes that were run in R-Studio for the two Bayesian models that were fitted for stunting and poverty separately and jointly.

A.1 R Codes for Bayesian GLMM Through MCMCs

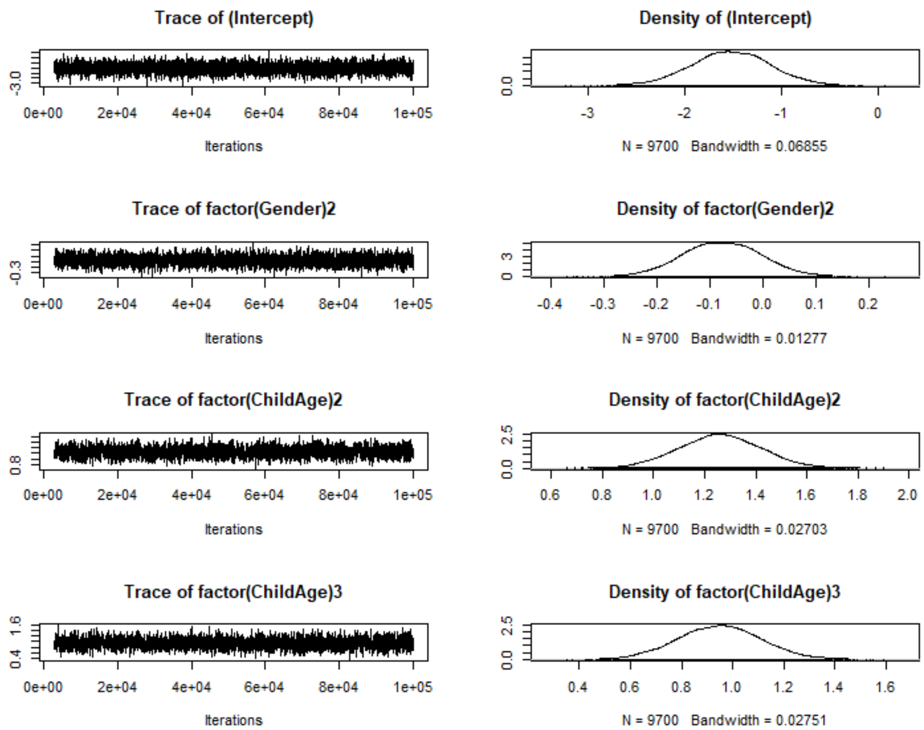
```
ModelStunted <- MCMCglmm(Stunted ~ 1 + factor(Gender) + factor(ChildAge) + factor(HeadshipGender)
+ factor(Region) + factor(MotherEduLevel) + factor(SourceWater) + factor(TypeRes) + fac-
tor(AnaemiaLevel) + BirthOrder + HeadshipAge + HouseholdSize + HHWealthIndexScore,
random = ~ HHNum:Cluster+Cluster, data = DATATHESIS22, family="categorical", prior=list(R=list(V=1,
fix=1),
G=list(G1=list(V=1, nu=1),G2=list(V=1, nu=1))),nitt=60000,slice=F,thin=10)

summary(ModelStunted)
plot(ModelStunted)

ModelPoverty <- MCMCglmm(Poverty ~ 1 + factor(HeadshipGender) + factor(Region) +
factor(MotherEduLevel) + factor(SourceWater) + factor(TypeRes) + HeadshipAge + House-
holdSize, random = ~ HHNum:Cluster+Cluster, data = DATATHESIS22, family="categorical",
prior=list(R=list(V=1, fix=1),
G=list(G1=list(V=1, nu=1),G2=list(V=1, nu=1))),nitt=60000,slice=F,thin=10)

summary(ModelPoverty)
plot(ModelPoverty)

joint_model2 <- MCMCglmm(Response ~ - 1 + Event+ Event:factor(Gender) + Event:factor(ChildAge)
+ Event:factor(HeadshipGender) + Event:factor(Region) + Event:factor(MotherEduLevel) +
Event:factor(SourceWater) + Event:factor(TypeRes) + Event:factor(AnaemiaLevel) + Event:BirthOrder
+ Event:HeadshipAge + Event:HouseholdSize, random = ~ Cluster,rcov = ~ us(Event):ID,
data = Joined_Data2, family = "categorical",prior= list(G = list(G1 = list(V = 1, nu = 1)),R =
list(V = diag(1), nu = 100)),nitt=60000,slice=F,thin=25, verbose = TRUE)

plot(joint_model2)
summary(joint_model1)
```

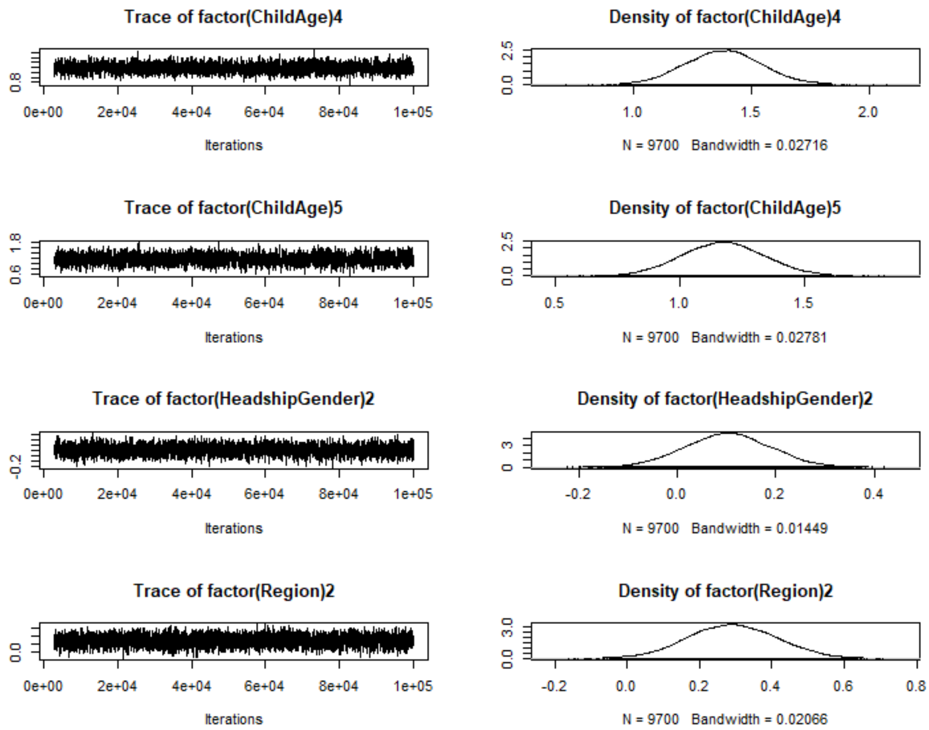
Appendix B

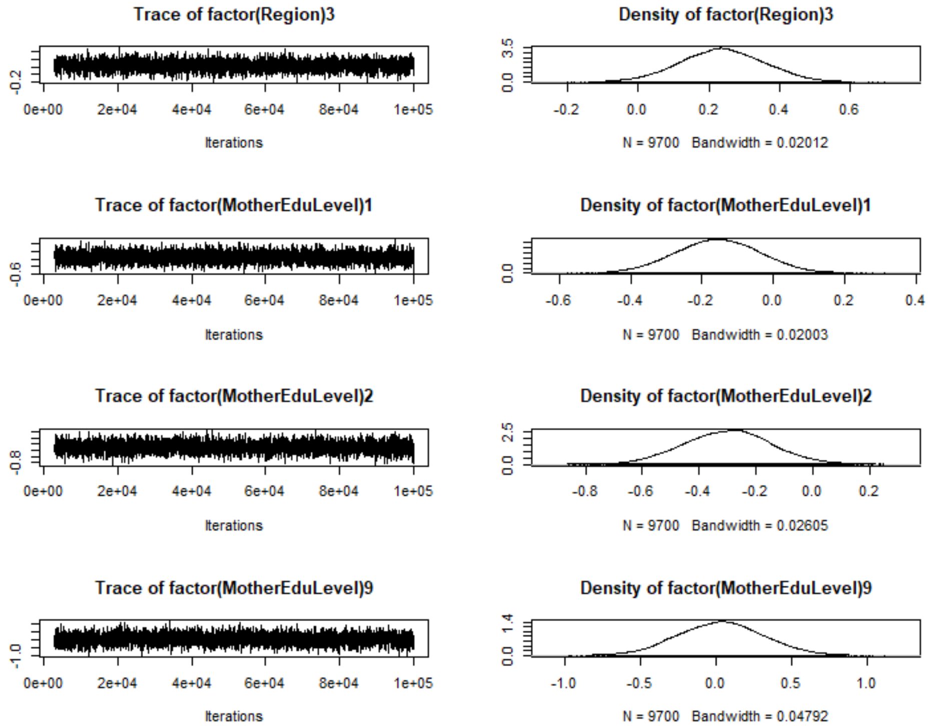
This section presents a sample of the trace plots for the two Bayesian models that were fitted for stunting and poverty separately.

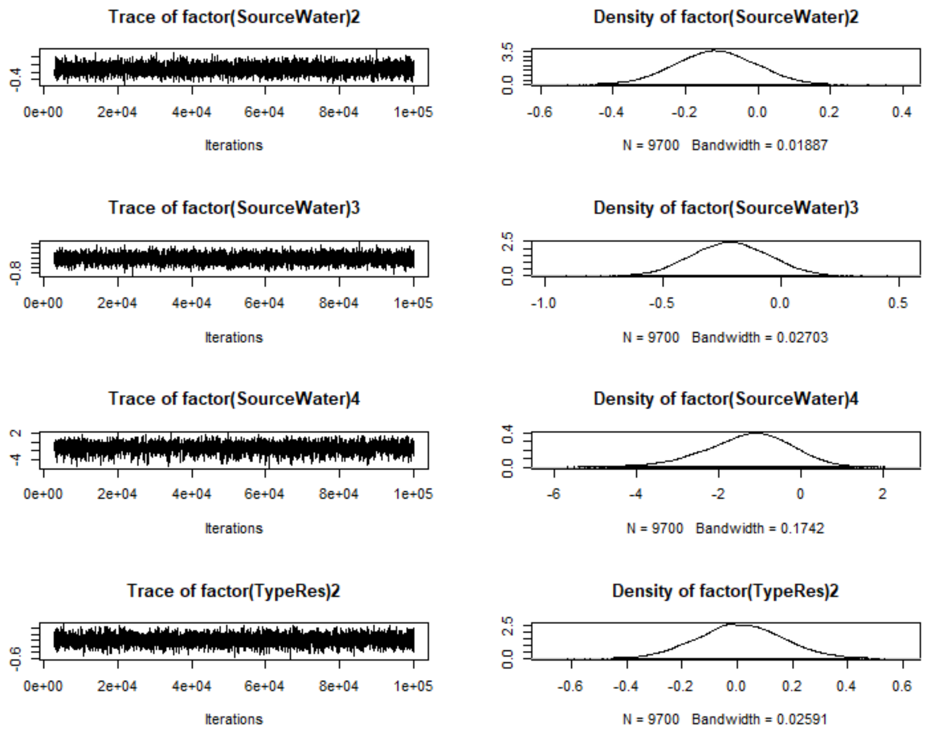
B.1 Bayesian GLMM trace plots for Stunting model

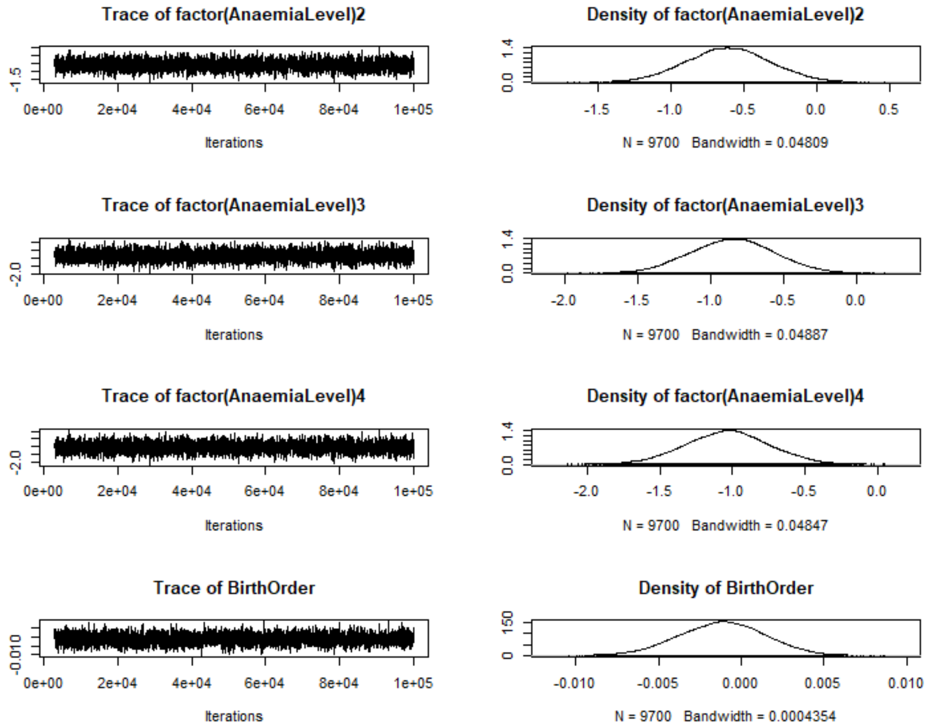
Presented below are the posterior distributions of the fixed and random effect trace plots for the stunting model, based on the analysis of 60000 iterations, 2000 burn in period, and with thin=10 in MCMCglmm.




Trace of factor(ChildAge)4
Iterations
Density of factor(ChildAge)4
N = 9700 Bandwidth = 0.02716
Trace of factor(ChildAge)5
Iterations
Density of factor(ChildAge)5
N = 9700 Bandwidth = 0.02781
Trace of factor(HeadshipGender)2
Iterations
Density of factor(HeadshipGender)2
N = 9700 Bandwidth = 0.01449
Trace of factor(Region)2
Iterations
Density of factor(Region)2
N = 9700 Bandwidth = 0.02066


Trace of factor(Region)3
Iterations
Density of factor(Region)3
N = 9700 Bandwidth = 0.02012
Trace of factor(MotherEduLevel)1
Iterations
Density of factor(MotherEduLevel)1
N = 9700 Bandwidth = 0.02003
Trace of factor(MotherEduLevel)2
Iterations
Density of factor(MotherEduLevel)2
N = 9700 Bandwidth = 0.02605
Trace of factor(MotherEduLevel)9
Iterations
Density of factor(MotherEduLevel)9
N = 9700 Bandwidth = 0.04792


Trace of factor(SourceWater)2
-0.4
0e+00 2e+04 4e+04 6e+04 8e+04 1e+05
Iterations
Density of factor(SourceWater)2
0.0 3.5
-0.6 -0.4 -0.2 0.0 0.2 0.4
N = 9700 Bandwidth = 0.01887
Trace of factor(SourceWater)3
-0.8
0e+00 2e+04 4e+04 6e+04 8e+04 1e+05
Iterations
Density of factor(SourceWater)3
0.0 2.5
-1.0 -0.5 0.0 0.5
N = 9700 Bandwidth = 0.02703
Trace of factor(SourceWater)4
-4 2
0e+00 2e+04 4e+04 6e+04 8e+04 1e+05
Iterations
Density of factor(SourceWater)4
0.0 0.4
-6 -4 -2 0 2
N = 9700 Bandwidth = 0.1742
Trace of factor(TypeRes)2
-0.6
0e+00 2e+04 4e+04 6e+04 8e+04 1e+05
Iterations
Density of factor(TypeRes)2
0.0 2.5
-0.6 -0.4 -0.2 0.0 0.2 0.4 0.6
N = 9700 Bandwidth = 0.02591


Trace of factor(AnaemiaLevel)2
-1.5
0e+00 2e+04 4e+04 6e+04 8e+04 1e+05
Iterations
Density of factor(AnaemiaLevel)2
0.0 1.4
-1.5 -1.0 -0.5 0.0 0.5
N = 9700 Bandwidth = 0.04809
Trace of factor(AnaemiaLevel)3
-2.0
0e+00 2e+04 4e+04 6e+04 8e+04 1e+05
Iterations
Density of factor(AnaemiaLevel)3
0.0 1.4
-2.0 -1.5 -1.0 -0.5 0.0
N = 9700 Bandwidth = 0.04887
Trace of factor(AnaemiaLevel)4
-2.0
0e+00 2e+04 4e+04 6e+04 8e+04 1e+05
Iterations
Density of factor(AnaemiaLevel)4
0.0 1.4
-2.0 -1.5 -1.0 -0.5 0.0
N = 9700 Bandwidth = 0.04847
Trace of BirthOrder
-0.010
0e+00 2e+04 4e+04 6e+04 8e+04 1e+05
Iterations
Density of BirthOrder
0 150
-0.010 -0.005 0.000 0.005 0.010
N = 9700 Bandwidth = 0.0004354

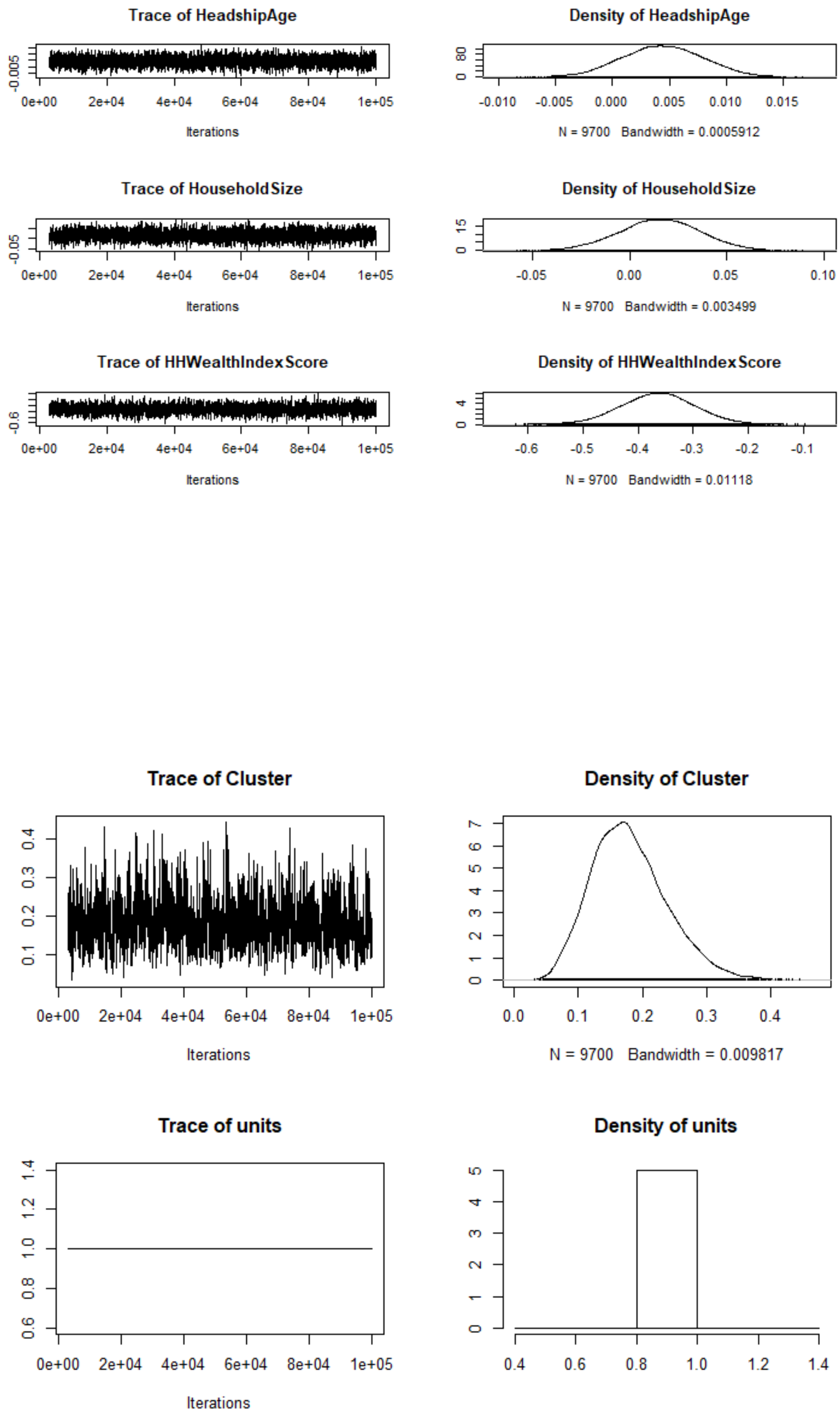
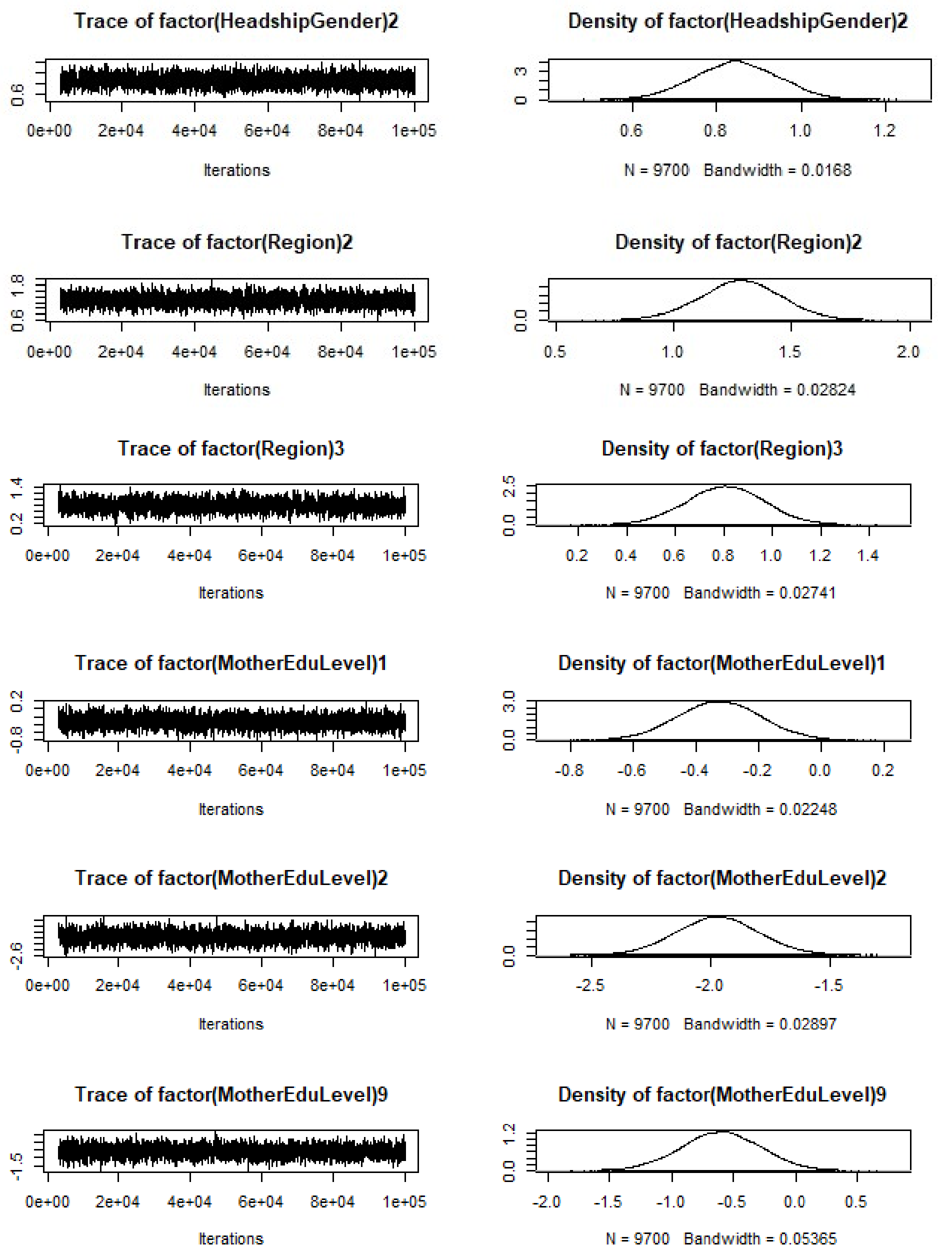
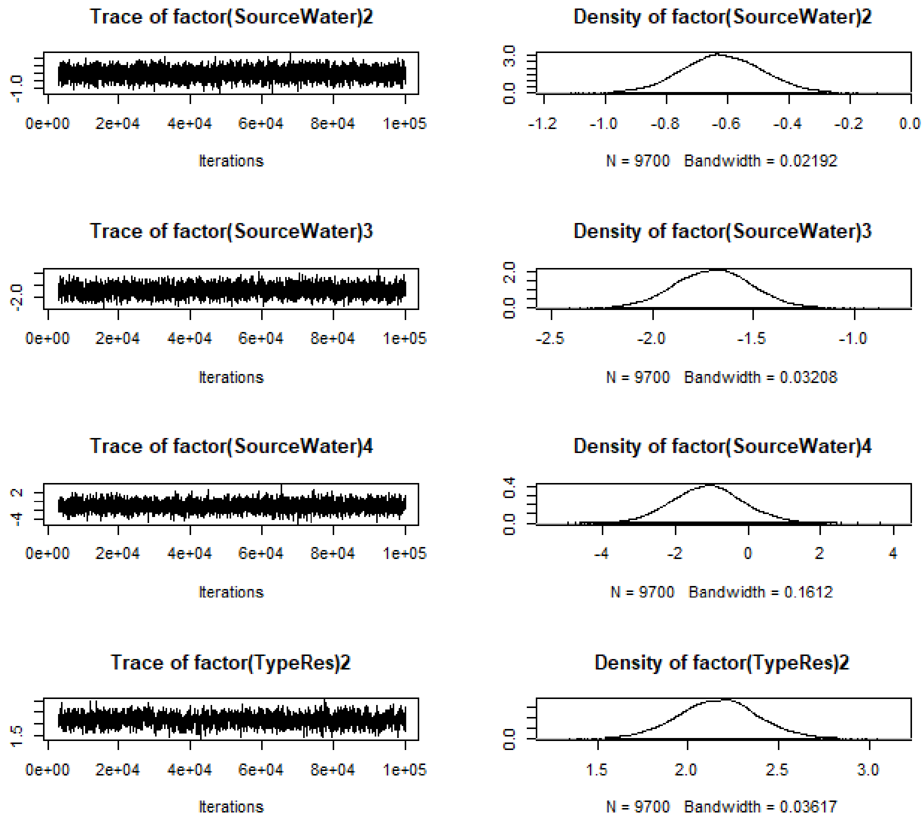


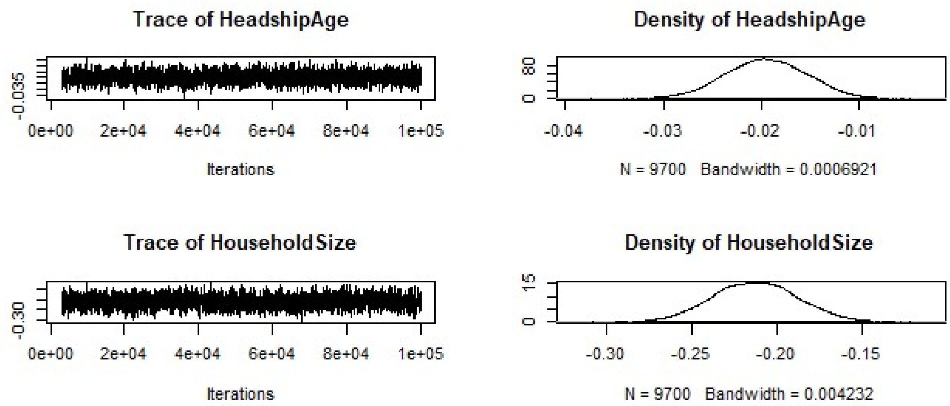
Figure 6.1: Trace Plots for Stunting model

B.2 Bayesian GLMM trace plots for the poverty model

Presented below are the posterior distribution of the fixed and random effect trace plots for the poverty model, based on the analysis of 60000 iterations, 2000 burn in period, and with thin=10 in MCMCglmm.




Trace of factor(SourceWater)2
Density of factor(SourceWater)2
N = 9700 Bandwidth = 0.02192
Trace of factor(SourceWater)3
Density of factor(SourceWater)3
N = 9700 Bandwidth = 0.03208
Trace of factor(SourceWater)4
Density of factor(SourceWater)4
N = 9700 Bandwidth = 0.1612
Trace of factor(TypeRes)2
Density of factor(TypeRes)2
N = 9700 Bandwidth = 0.03617
Iterations


Trace of HeadshipAge
Density of HeadshipAge
N = 9700 Bandwidth = 0.0006921
Trace of HouseholdSize
Density of HouseholdSize
N = 9700 Bandwidth = 0.004232
Iterations

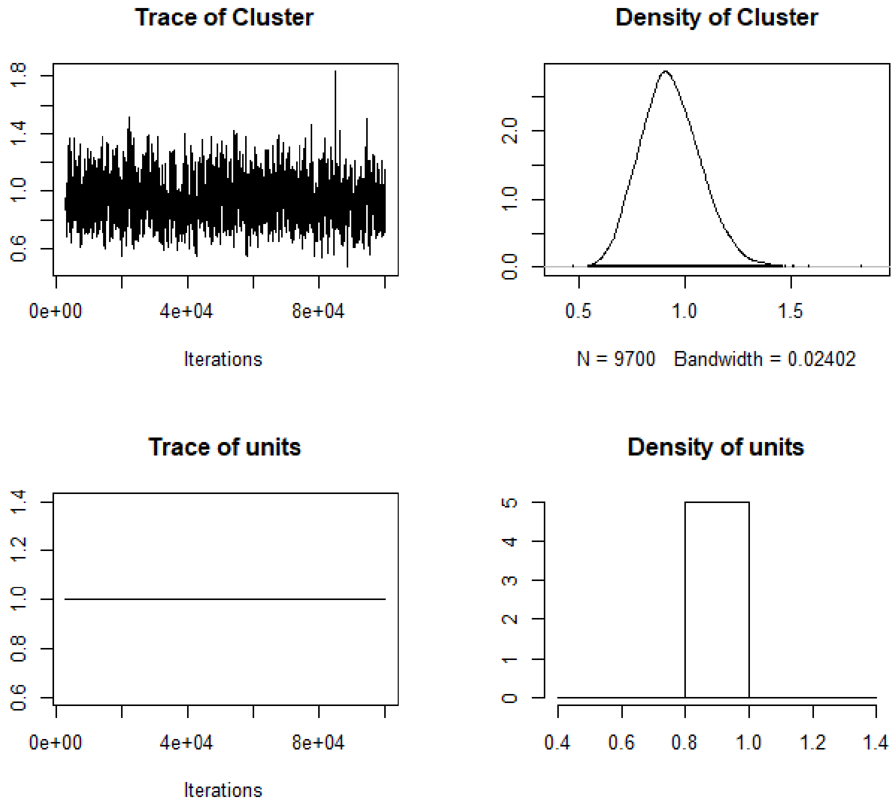
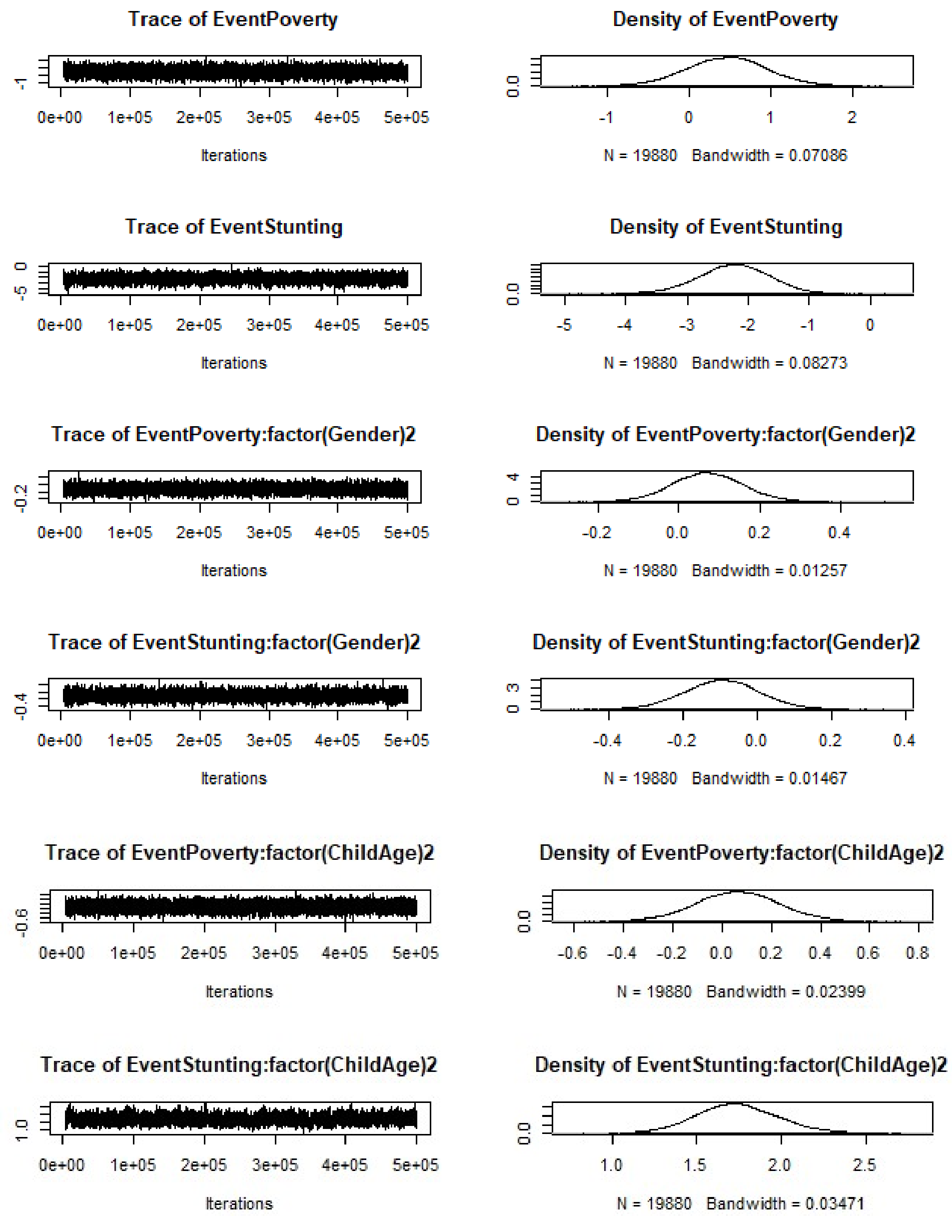
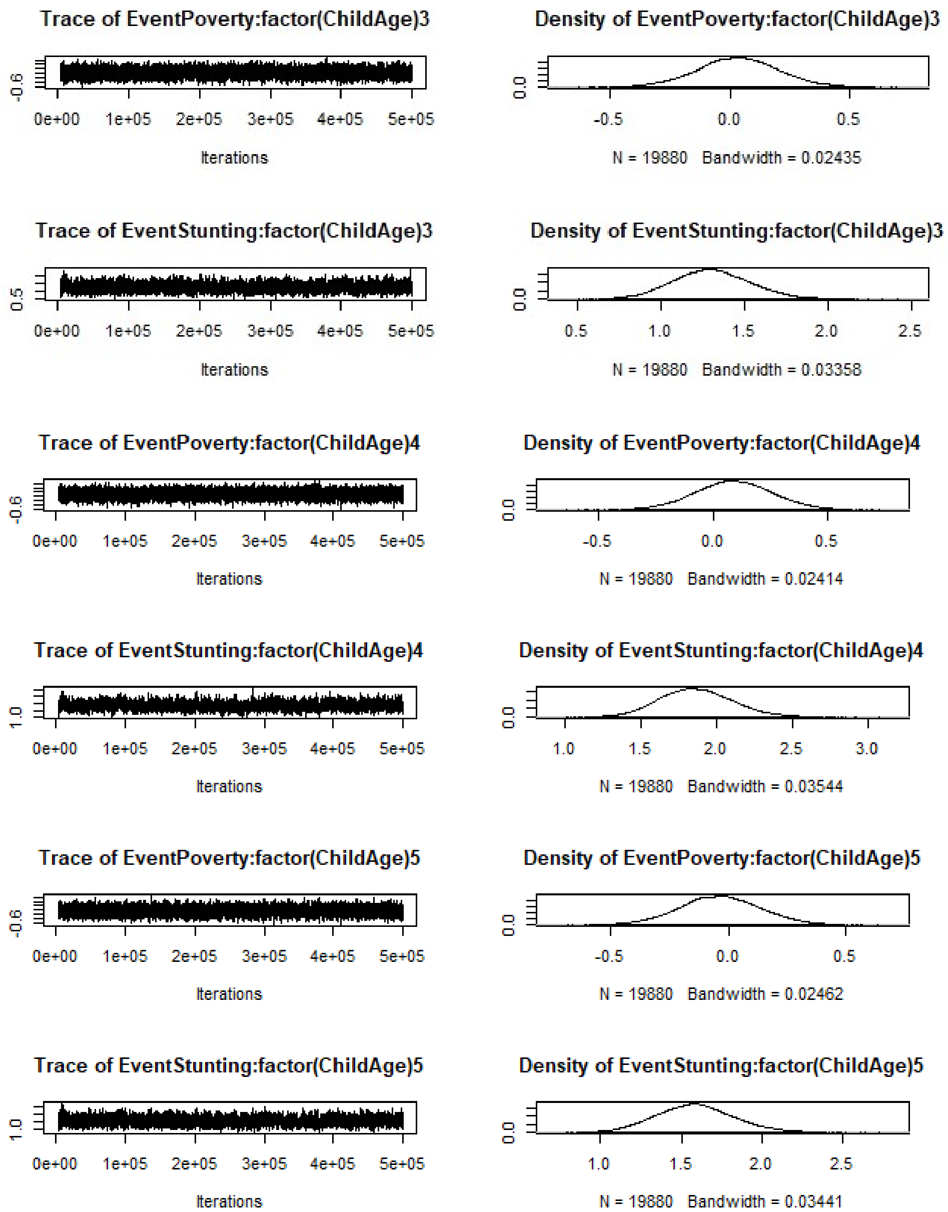


Figure 6.2: Trace Plots for Household Poverty model

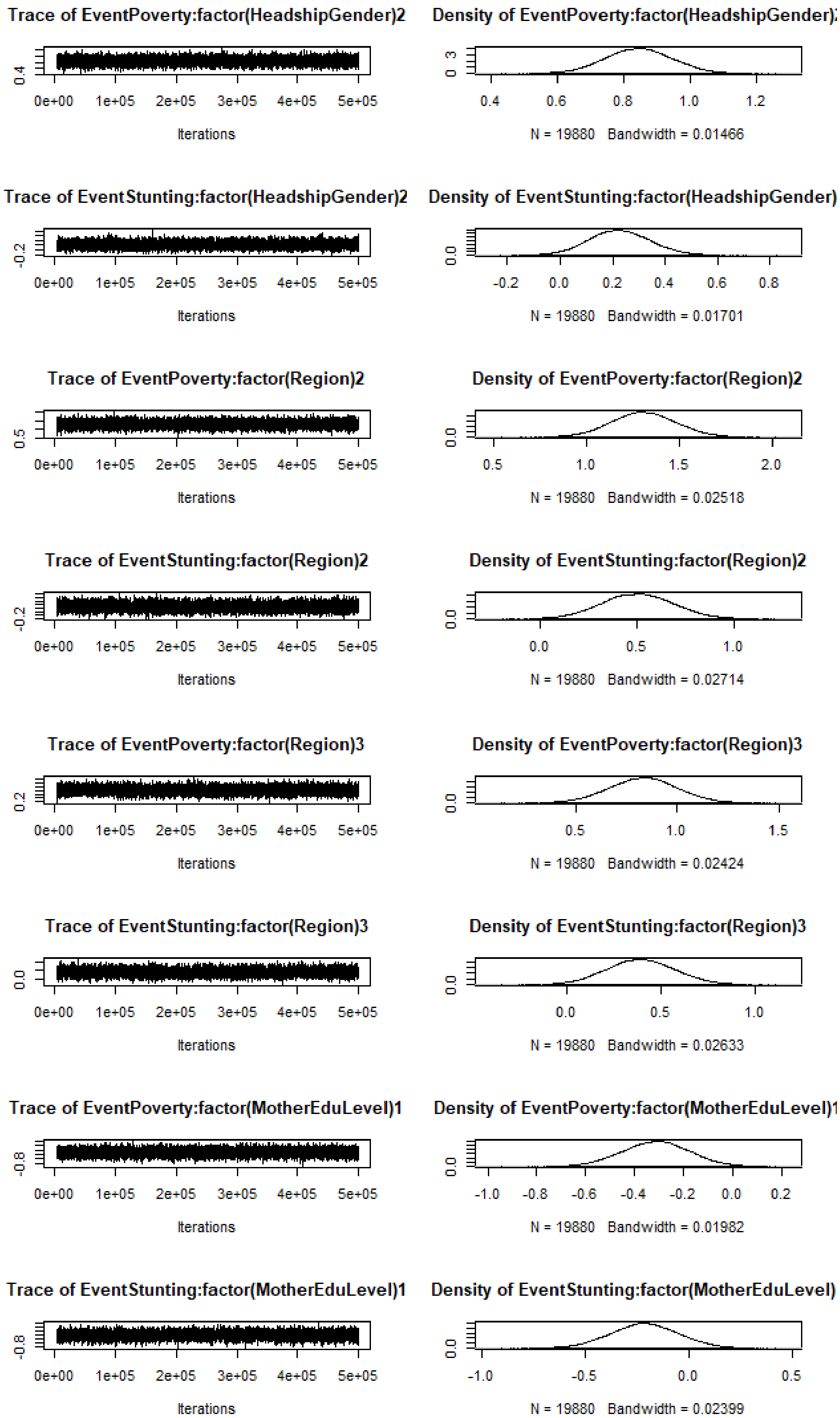
B.2 Bayesian GLMM trace plots for the joint model of poverty and malnutrition

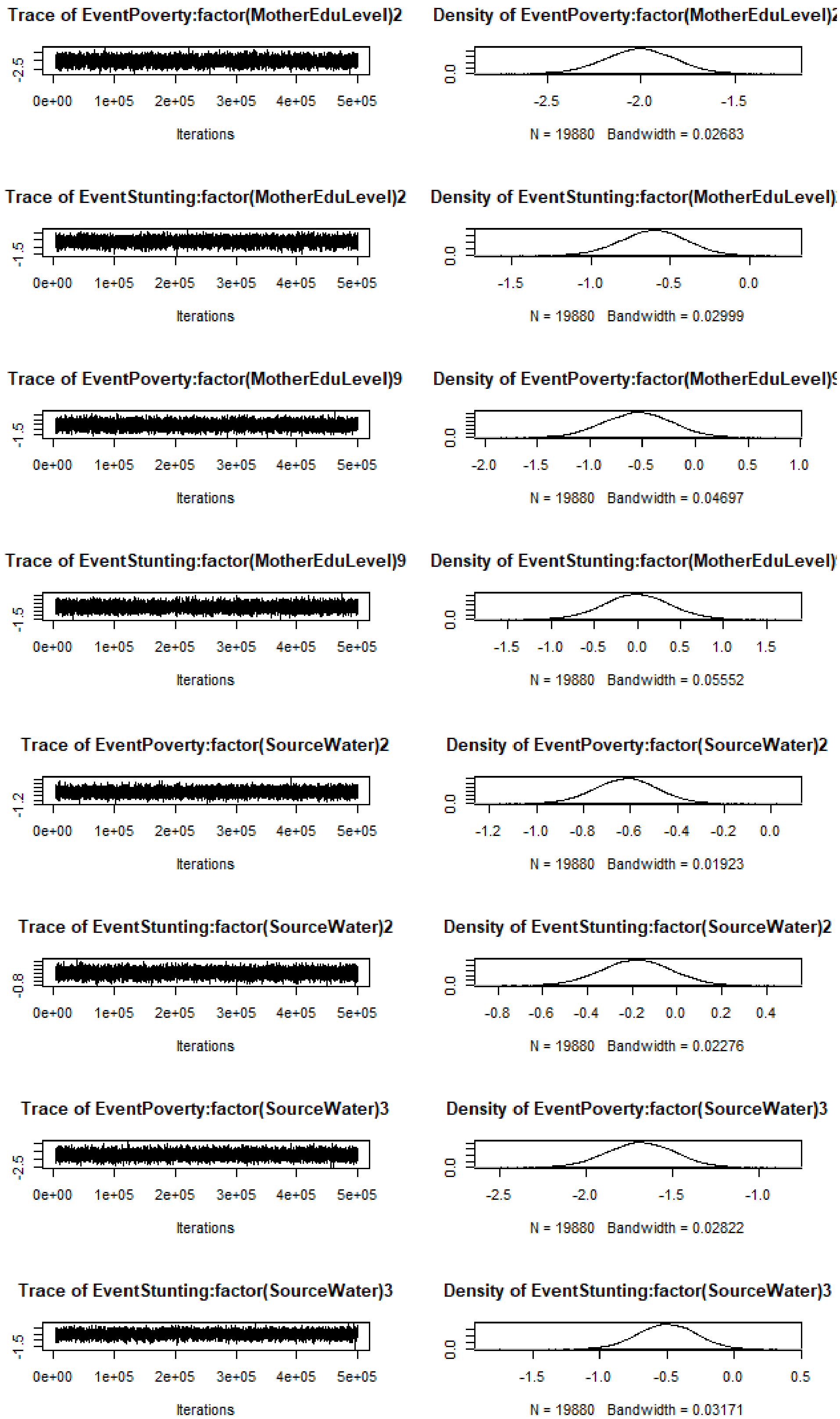
Presented below are some of the posterior distribution of the fixed and random effect trace plots for the joint model of malnutrition and poverty, based on the analysis of 60000 iterations, 2000 burn in period, and with thin=25 in MCMCglmm.

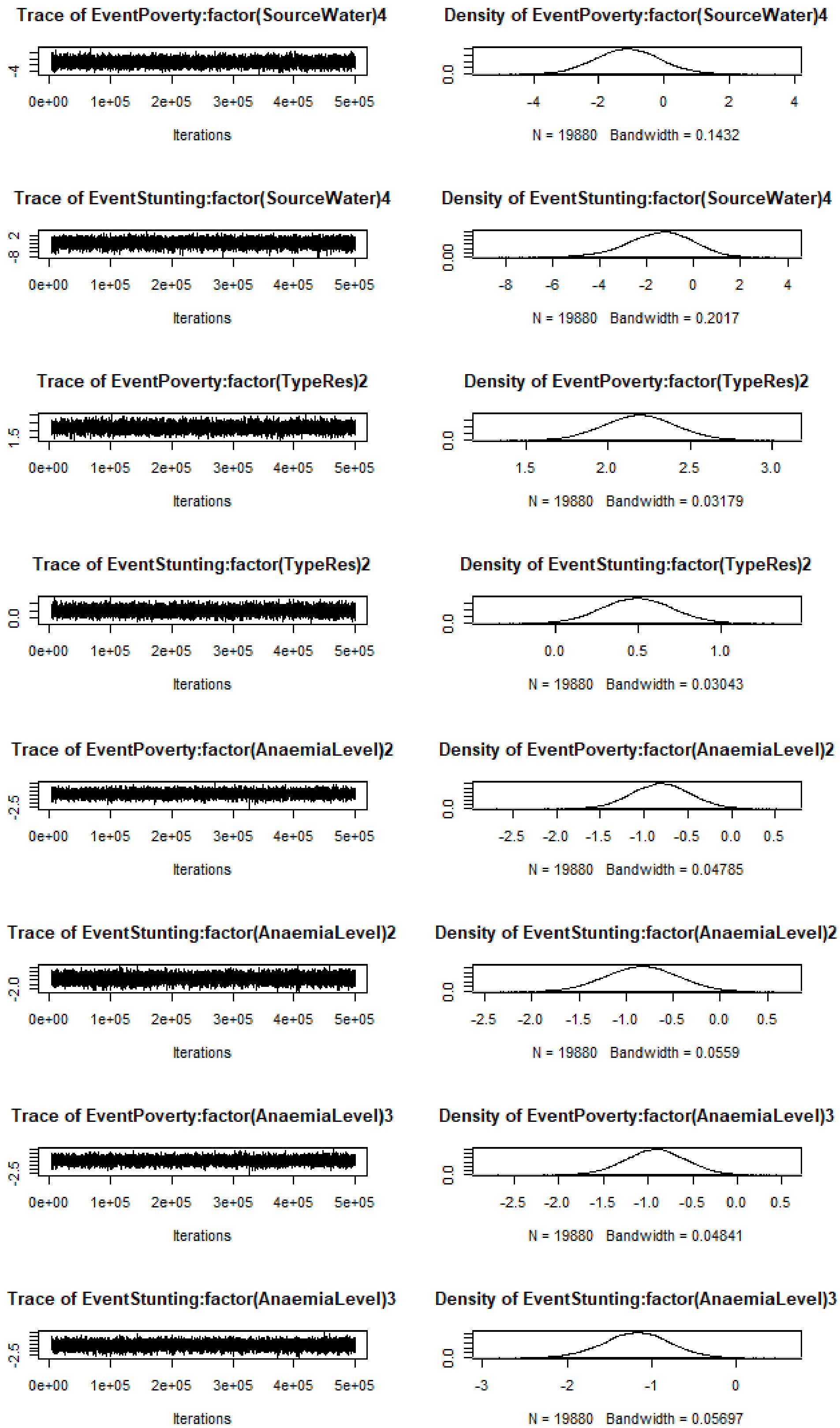


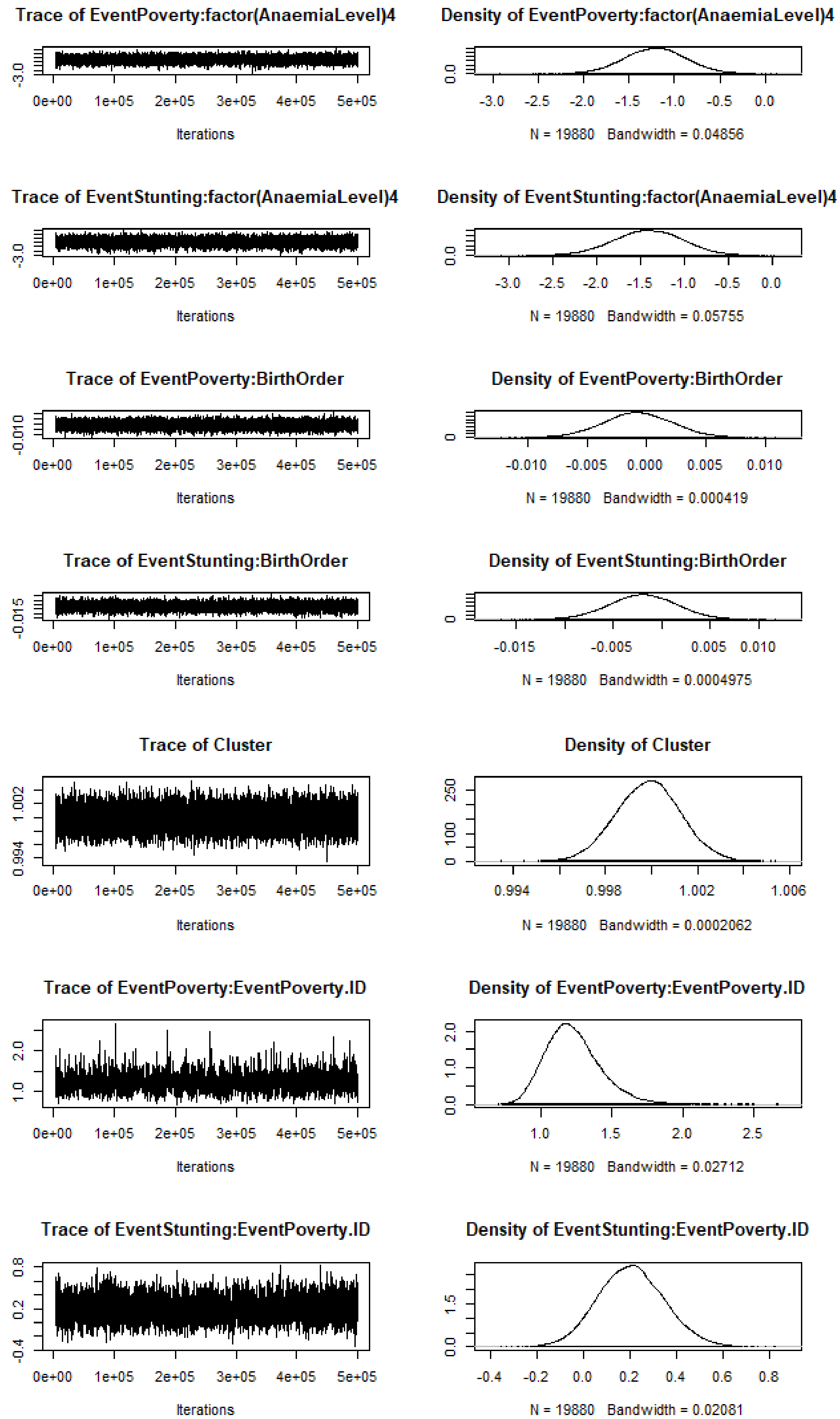


Trace of EventPoverty:factor(ChildAge)3
-0.6
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventPoverty:factor(ChildAge)3
0.0
-0.5 0.0 0.5
N = 19880 Bandwidth = 0.02435
Trace of EventStunting:factor(ChildAge)3
0.5
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventStunting:factor(ChildAge)3
0.0
0.5 1.0 1.5 2.0 2.5
N = 19880 Bandwidth = 0.03358
Trace of EventPoverty:factor(ChildAge)4
-0.6
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventPoverty:factor(ChildAge)4
0.0
-0.5 0.0 0.5
N = 19880 Bandwidth = 0.02414
Trace of EventStunting:factor(ChildAge)4
1.0
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventStunting:factor(ChildAge)4
0.0
1.0 1.5 2.0 2.5 3.0
N = 19880 Bandwidth = 0.03544
Trace of EventPoverty:factor(ChildAge)5
-0.6
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventPoverty:factor(ChildAge)5
0.0
-0.5 0.0 0.5
N = 19880 Bandwidth = 0.02462
Trace of EventStunting:factor(ChildAge)5
1.0
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventStunting:factor(ChildAge)5
0.0
1.0 1.5 2.0 2.5
N = 19880 Bandwidth = 0.03441


Trace of EventPoverty:factor(HeadshipGender)2
0.4
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventPoverty:factor(HeadshipGender)
0 3
0.4 0.6 0.8 1.0 1.2
N = 19880 Bandwidth = 0.01466
Trace of EventStunting:factor(HeadshipGender)2
-0.2
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventStunting:factor(HeadshipGender)
0.0
-0.2 0.0 0.2 0.4 0.6 0.8
N = 19880 Bandwidth = 0.01701
Trace of EventPoverty:factor(Region)2
0.5
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventPoverty:factor(Region)2
0.0
0.5 1.0 1.5 2.0
N = 19880 Bandwidth = 0.02518
Trace of EventStunting:factor(Region)2
-0.2
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventStunting:factor(Region)2
0.0
0.0 0.5 1.0
N = 19880 Bandwidth = 0.02714
Trace of EventPoverty:factor(Region)3
0.2
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventPoverty:factor(Region)3
0.0
0.5 1.0 1.5
N = 19880 Bandwidth = 0.02424
Trace of EventStunting:factor(Region)3
0.0
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventStunting:factor(Region)3
0.0
0.0 0.5 1.0
N = 19880 Bandwidth = 0.02633
Trace of EventPoverty:factor(MotherEduLevel)1
-0.8
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventPoverty:factor(MotherEduLevel)1
0.0
-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2
N = 19880 Bandwidth = 0.01982
Trace of EventStunting:factor(MotherEduLevel)1
-0.8
0e+00 1e+05 2e+05 3e+05 4e+05 5e+05
Iterations
Density of EventStunting:factor(MotherEduLevel)
0.0
-1.0 -0.5 0.0 0.5
N = 19880 Bandwidth = 0.02399


Trace of EventPoverty:factor(MotherEduLevel)2
Density of EventPoverty:factor(MotherEduLevel)2
Iterations
N = 19880 Bandwidth = 0.02683
Trace of EventStunting:factor(MotherEduLevel)2
Density of EventStunting:factor(MotherEduLevel)
Iterations
N = 19880 Bandwidth = 0.02999
Trace of EventPoverty:factor(MotherEduLevel)9
Density of EventPoverty:factor(MotherEduLevel)
Iterations
N = 19880 Bandwidth = 0.04697
Trace of EventStunting:factor(MotherEduLevel)9
Density of EventStunting:factor(MotherEduLevel)
Iterations
N = 19880 Bandwidth = 0.05552
Trace of EventPoverty:factor(SourceWater)2
Density of EventPoverty:factor(SourceWater)2
Iterations
N = 19880 Bandwidth = 0.01923
Trace of EventStunting:factor(SourceWater)2
Density of EventStunting:factor(SourceWater)2
Iterations
N = 19880 Bandwidth = 0.02276
Trace of EventPoverty:factor(SourceWater)3
Density of EventPoverty:factor(SourceWater)3
Iterations
N = 19880 Bandwidth = 0.02822
Trace of EventStunting:factor(SourceWater)3
Density of EventStunting:factor(SourceWater)3
Iterations
N = 19880 Bandwidth = 0.03171


Trace of EventPoverty:factor(SourceWater)4
Iterations
Density of EventPoverty:factor(SourceWater)4
N = 19880 Bandwidth = 0.1432
Trace of EventStunting:factor(SourceWater)4
Iterations
Density of EventStunting:factor(SourceWater)4
N = 19880 Bandwidth = 0.2017
Trace of EventPoverty:factor(TypeRes)2
Iterations
Density of EventPoverty:factor(TypeRes)2
N = 19880 Bandwidth = 0.03179
Trace of EventStunting:factor(TypeRes)2
Iterations
Density of EventStunting:factor(TypeRes)2
N = 19880 Bandwidth = 0.03043
Trace of EventPoverty:factor(AnaemiaLevel)2
Iterations
Density of EventPoverty:factor(AnaemiaLevel)2
N = 19880 Bandwidth = 0.04785
Trace of EventStunting:factor(AnaemiaLevel)2
Iterations
Density of EventStunting:factor(AnaemiaLevel)2
N = 19880 Bandwidth = 0.0559
Trace of EventPoverty:factor(AnaemiaLevel)3
Iterations
Density of EventPoverty:factor(AnaemiaLevel)3
N = 19880 Bandwidth = 0.04841
Trace of EventStunting:factor(AnaemiaLevel)3
Iterations
Density of EventStunting:factor(AnaemiaLevel)3
N = 19880 Bandwidth = 0.05697


Trace of EventPoverty:factor(AnaemiaLevel)4
Iterations
Density of EventPoverty:factor(AnaemiaLevel)4
N = 19880 Bandwidth = 0.04856
Trace of EventStunting:factor(AnaemiaLevel)4
Iterations
Density of EventStunting:factor(AnaemiaLevel)4
N = 19880 Bandwidth = 0.05755
Trace of EventPoverty:BirthOrder
Iterations
Density of EventPoverty:BirthOrder
N = 19880 Bandwidth = 0.000419
Trace of EventStunting:BirthOrder
Iterations
Density of EventStunting:BirthOrder
N = 19880 Bandwidth = 0.0004975
Trace of Cluster
Iterations
Density of Cluster
N = 19880 Bandwidth = 0.0002062
Trace of EventPoverty:EventPoverty.ID
Iterations
Density of EventPoverty:EventPoverty.ID
N = 19880 Bandwidth = 0.02712
Trace of EventStunting:EventPoverty.ID
Iterations
Density of EventStunting:EventPoverty.ID
N = 19880 Bandwidth = 0.02081

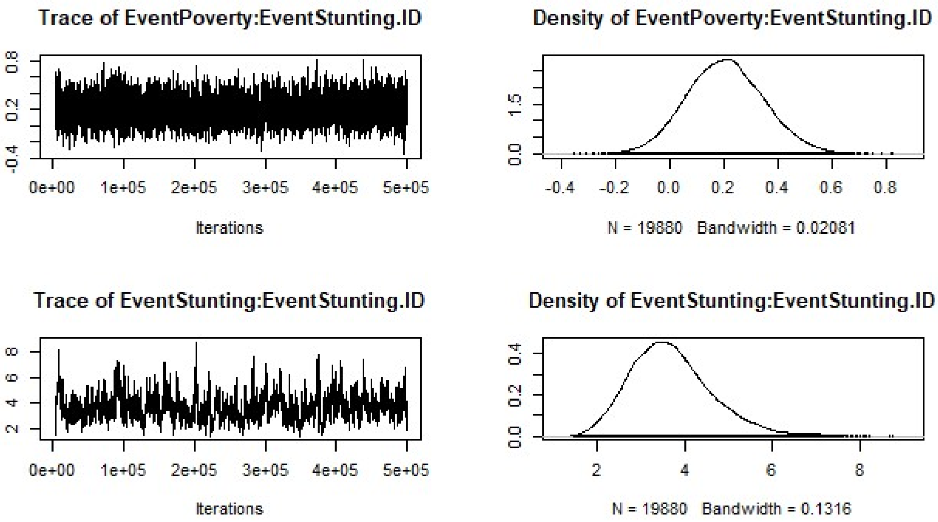


Figure 6.3: Trace Plots for The Joint Model of Poverty And Malnutrition