UNIVERSITY OF CAPE COAST

ESTIMATION TECHNIQUES IN GENERALIZED LINEAR MIXED MODELS WITH APPLICATION TO DISEASE IMPACT MODELLING

DAVID YAO MENSAH

2015

UNIVERSITY OF CAPE COAST

ESTIMATION TECHNIQUES IN GENERALIZED LINEAR MIXED MODELS WITH APPLICATION TO DISEASE IMPACT MODELLING

BY

DAVID YAO MENSAH

Thesis submitted to the Department of Mathematics and Statistics of the School of Physical Sciences, College of Agriculture and Natural Sciences, University of Cape Coast, in partial fulfilment of the requirements for the award of Doctor of Philosophy degree in Statistics

JULY 2015

DECLARATION

CANDIDATE'S DECLARATION

I hereby declare that this thesis is the result of my own original research and that no part of it has been presented for another degree in this university or elsewhere

Candidate's Signature:……………………………………. Date:………………

Name:…………………………………………………...

SUPERVISORS' DECLARATION

We hereby declare that the preparation and presentation of the thesis were supervised in accordance with the guidelines on supervision of thesis laid down by the University of Cape Coast.

Principal Supervisor's Signature:…………………………… Date:……………

Name:………………………………………………......

Co-Supervisor's Signature:…………………………………. Date:……………

Name:……………………………………………….....

ABSTRACT

This study applies the theory of Generalised linear Mixed Models (GLMMs) to survey data on disease impact in Ghana. It determines the variables that are responsible for making dependent members of households feel the impact of illness and/or death for three identified types of households. It assesses four models in terms of these variables generated using the Maximum Mean Pseudo-Likelihood (MMPL) and the Residual Mean Pseudo-Likelihood (RMPL) techniques in SAS. For all four models considered, the MMPL produces more suitable models than the RMPL. The impact of illness and/or death on HIV/AIDS, Other Illness/Deaths or No Illness/Death households is felt in the areas of reallocation of dependents' time, dependents having to work harder to substitute for lost income, dependents leaving work to care for the sick, and household reducing expenditure as a result of illness and/or death. It is found that the degree of impact depends on marital status, sex or tribe of household headship, remoteness of occurrence of mortality/morbidity, total asset value, and level of annual adult's health expenditure. It is also found that in almost all cases examined, the HIV/AIDS household suffers a significant impact compared to No/Other disease household. The findings indicate that there should be a continued effort at reducing not only the incidence but also the impact of HIV/AIDS on households.

ACKNOWLEDGEMENTS

I wish to thank Professor Nicholas N. N. Nsowah-Nuamah and Dr. Bismark Nkansah for their willingness to be my supervisors, for guiding me through my thesis and for providing me with valuable suggestions. I would also like to gratefully acknowledge the excellent guidance and constant encouragement provided by Professor Benony Gordor during the past few years to get this work completed.

I would like to specially thank Professor Awuku Sakyi-Amoa formerly of the Ghana Aids Commission (GAC) for his unfailing assistance in readily linking me to NAP+ to collect data for this thesis. I thank the national executive of NAP+ (2008) for their unflinching support during my data collection.

I have to say 'thank-you' to Victor Mensah and Jerry Adjei Akuffo who collected part of the data, all my friends and family for their moral support, wherever they are.

I finally thank the former Head of the Department of Mathematics and Statistics of the University of Cape Coast, Prof. (Mrs.) N. Mensah, the current Head of Department, Dr. Henry Amankwah, staff of the department and my course mates for always being there for me.

DEDICATION

To the memory of Victoria Yawo Mensah and the late Isaac Ayodele Komla Mensah

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

Background

Since the 1980s, when the first cases of AIDS were identified, HIV/AIDS has emerged as one of the leading challenges to global public health. Particularly in sub-Saharan Africa, where the majority of HIV and AIDS cases are concentrated, the epidemic continues to take an extraordinary human toll.

Given the need to understand better the levels and trends of the HIV epidemic, and the limited information on which to base these estimates, mathematical models can make a valuable contribution. The goal of such models, like any modelling exercise, is to extract as much information as possible from the available data and provide an accurate representation of both the knowledge and uncertainty about the epidemic (Kareem et al., 2010).

Mathematical models have long provided basic insights for HIV/AIDS control. The recent success of the Onchocerciasis Control Program in West Africa shows that models can make great pragmatic contributions to intervention programs if the modelling is integrated into the overall program, and if the participants are clear about what models can and cannot do (Hopkins et al., 2005).

There needs to be realistic expectations regarding the kind of output models can generate. In particular, in this context, models cannot provide accurate numerical predictions of outcomes; they can be used to forecast, but only in fairly gross terms. The biological, social, and other systems involved are sufficiently complex that it may not be possible to even define all of the variables, much less get precise predictions about their interactions and overall results in a specific real-world situation. Thus, the key is to look for large differences between different models, and between different interventions in the same modeling scenario. That is; mathematical models can be used to 1) systematically compare alternate strategies, 2) determine the key issues in decision-making, and 3) identify gaps in current knowledge. Mathematical models can help us figure out which decisions will have the largest impacts on outcomes and can provide comprehensive examinations of the assumptions that feed into decisions in a way that purely verbal reasoning and debate cannot (Mckenzie and Samba, 2004).

Studies conducted in stable transmission areas, such as Khan, (1966), Ettling and Shepard, (1991), and Attanayake et al. (2000) have established that HIV/AIDS causes substantial losses to households in the form of foregone income, treatment costs, missed schooling, and decreased agricultural production.

The most important level at which to measure impact is at the household and community level (Savigny and Binka, 2004; Agyepong et al., 2004). Despite imprecision, the HIV/AIDS toll has been relatively well-quantified clinically and epidemiologically. In economic and social terms, it is less well understood. Estimates have suggested that HIV/AIDS costs African countries about $12 billion annually and may considerably retard economic development (Gallup and Sachs, 2001). An African family may spend up to 25% of its income on HIV/AIDS prevention and control (The Abuja Declaration and the Plan of Action, April 25, 2000). The economic burden of ill health on individual households can be substantial and in some cases catastrophic, especially for poor households. Russell (2004) reviews studies that have measured the economic costs and consequences of illness for households, focusing on HIV/AIDS and tuberculosis. He finds that illness imposes high and regressive cost burdens on patients and their families in poor settings.

The explosive nature of the HIV/AIDS epidemics also seems to overwhelm the social and administrative infrastructure that would otherwise exist to cope with them. Even the orderly scheduling of funerals might be disrupted by high case fatality rates, causing great stress to families who lose family members. The added uncertainty and explosive intensity of the HIV/AIDS epidemic may affect not only the magnitude but also the nature of economic burdens imposed compared with regions where HIV/AIDS transmission is more stable. Quantitative estimates of the economic effects of HIV/AIDS are presently lacking (Kiszewski and Teklehaimanot, 2004).

General Impact of the Study

Studies conducted indicate dramatic changes in the household structure and composition as a result of HIV/AIDS, which has a bearing on agricultural production (Barnett and Haslwimmer, 1995). The economic and social consequences of the disease directly affect the family (Barnett et al., 2001). In the absence of functioning medical care systems in African countries, medical costs and caring for sick family members must be borne entirely by the nuclear family or by the extended family network. In addition to the medical costs, which include the cost of drugs and traditional medical treatment, funeral expenses of family members are a heavy burden on the family budget. Funeral costs are one area of burden to families as they appear to be even higher than medical expenses in some settings. All said and done, it has been observed that the decline in farm income caused by a decline in farm activities due to incapacitation and livestock production, coupled with an increase in medical expenses and funeral costs, can lead to the breakdown of the nuclear family and the traditional support system (Agartha et al., 2010). The inter-linkages between the increase of HIV/AIDS-related mortality and morbidity, the lack of farm inputs and labour force, the deterioration of household economy and the impact on education, health and the social system, which eventually lead to a breakdown of the traditional coping mechanisms, are enormous (Ugwuanya, 2003).

Because HIV/AIDS infects mainly adults during their sexually active years and is inevitably fatal, the socioeconomic implications of HIV/AIDS for development are immense (Niehof, 2004). At the family level, the death of an adult during his or her sexually active years means the loss of a family member of prime working age whose foregone income can adversely affect the welfare of surviving family members, especially if the deceased is also the family's main breadwinner (Loevinsohn and Gillespie, 2003). This impact will be even worse if the family is a low-income family, because such families generally possess few resources, and are thus less able to cope with increased medical care costs and other related expenses, in addition to the foregone earnings of the ill family member. Hence, HIV/AIDS not only increases mortality, but also immiserizes the poor and widens income inequality between the haves and the have nots (Loevinsohn and Gillespie, 2003).

Evidence from a similar study by Pitayanon et al. (1997) suggests that AIDS interventions can no longer focus primarily on the infected individual and ways of preventing additional infection, but must also address the growing needs of those who are affected but uninfected, that is, the family, friends, and the whole community. This is because we now know that the epidemic's toll will be measured not only in terms of lives lost, but in the progressive circle of reduced functioning rippling through families, communities, and regions. This will be reflected not only in lost economic productivity, but in increasing social burdens, such as caring for children orphaned by the epidemic.

Farmers have developed mechanisms to cope with the impacts of HIV/AIDS on their rural livelihood strategies (Kwaramba, 1997). Traditionally, in emergency situations caused by natural disasters and in situations of hardship, the extended family network has developed successful coping mechanisms, which are still operational in pre-impact and early impact communities.

Traditional coping mechanisms are based mainly on returns to labour at the farm and/or family unit. Even the contribution of child labour may be increased (with children, particularly girls, withdrawn from school) as the family struggles to maintain the current cropping patterns. But, as a family becomes more impoverished, it may have little choice but to produce for its own consumption needs. Even then, family nutrition levels could be gradually compromised. It is not uncommon in full-impact districts/communities to observe entire families of children with elderly grandparents as their only form of support.

Since HIV/AIDS is above all a sexually transmitted disease, very often at least one family member is affected and dies. As a result, the entire assets and savings of many families, which are generally meagre before the onset of the disease, may be completely depleted in the bid to cure the ill family member, leaving the surviving family members without any means of support. A study in Uganda has shown that the burden of the socio-economic impact of HIV/AIDS is disproportionately affecting rural women (Barnett and Blaikie, 1992a). In the districts studied, more households were found to be headed by AIDS widows than by AIDS widowers. Widows with dependent children became entrenched in poverty as a result of the socio-economic pressures related to HIV/AIDS. Widows lost access to land, labour, inputs, credit and support services. HIV/AIDS stigmatization compounded their situation further, as assistance from the extended family and the community, their main safety net, was severed. The extent to which malnutrition rates in affected households rise depends on the type of coping mechanisms, household resource constraints, socio-cultural context and emotional stress. As the ability to produce and accumulate food and income decreases, the household falls into a downward spiral of increasing dependency ratios, poorer nutrition and health, increasing expenditure of resources (time and money) on health problems, more food shortages, decreasing household viability and increasing reliance on support from extended family and the wider community. The effects of HIV/AIDS on rural households, the likely impact of the disease on farmers' health and the nutrition of farm families manifest in different ways and over time.

Statement of the Problem

In Ghana, despite the setting up of a ministry for formal social welfare systems that cushion people in peculiar situations, e.g. the unemployed, sick, disabled etc., means that the dependent population (i.e. persons under fifteen years of age and persons above sixty years of age) is largely catered for by the working members of their households or families. However, the working population is the very population which is reported to have been hardest hit by the HIV/AIDS pandemic.

Available statistics (GNA, 2009) indicate that the HIV/AIDS situation in Ghana calls for concern and the need for precise studies to assess the impact it has on the dependent population, and policy efforts put in place to stem such impacts. However, no research work appears to have been done in Ghana over the years in employing statistical models in estimating the impact of HIV/AIDS morbidity and mortality on the dependent population. However, some amount of work in this regard has been attempted elsewhere in Thailand (Pitayanon et al., 1997). The Thailand study tried to identify some of the relative impacts that HIV/AIDS has had on the dependent members of the typical Thai household, and tried to compare the relative impact of HIV/AIDS on the household to the impact of other diseases on the household.

This study is a build-up on what was done by Pitayanon et al. (1997). Their study was based on only mortality. However, this study is motivated by the fact that households are affected economically not only by HIV/AIDS mortality but by HIV/AIDS morbidity as well. Mortality alone is not enough to present the full picture of the impact of HIV/AIDS on the household, and hence the dependent population. Pitayanon et al. (1997) were about the first to model the impact of HIV/AIDS on the household. In their study, they attempted running a binary logistic regression, among others, with which to estimate the economic impact of HIV/AIDS on the household. However, their study was constrained by inadequacy of data. This study addresses the data inadequacy problem by increasing the sample size beyond what was covered by Pitayanon et al. (1997).

The effect of addressing the data inadequacy problem offers first-hand information to policy makers on how to support affected households and what policies would be most appropriate in this regard. It also serves as an initial work to provoke further research into the modeling of different forms of impact that HIV/AIDS, as well as other diseases/illnesses, has on the household and different segments of the population.

Over the past few decades, several indicators have been developed to adjust mortality to reflect the impact of morbidity or disability. These measures fall into two basic categories, health expectancies and health gaps (Murray and Lopez, 2004; Jankovic et al., 2007; Jankovic, 2005). A member of the categories under health gaps, namely the disability-adjusted life years (DALYs), and related epidemiologic models, have been used to assess the morbidity and mortality at the locality, national and global levels (Melse et al., 2000; Mathers et al., 2001; Michaud et al., 2006; Kominski et al., 2002). These epidemiologic models, combined with demographic indicators, have been used in cost-effectiveness analyses of public health interventions (Schackman et al., 2007; Sinha et al., 2007; Llanos et al., 2007). Besides their inability to simultaneously determine the factors responsible for the morbidity impacts, these epidemiologic models are also not able to establish the fact that morbidity and mortality impacts vary across countries and socioeconomic strata, not making room for equity but rather running counter to it, thereby invalidating the use of YLD estimates as measures of disability (Grosse et al., 2009; Anand and Hanson, 1997). Moreover, these existing disease impact measurement models do not also consider impact measurement at the household level, though it serves as the most important level at which to measure impact (Savigny and Binka, 2004; Agyepong et al. (2000).

The motivation for this study therefore emanates from the work done by Pitayanon et al (1997), as well as the classical piece by Mzolo et al. (2009), Mzolo (2011) and Zhang et al. (2012). In these research works, there is a modelling gap with respect to the use of GLMMs in ascertaining the determinants of disease impact. This work is meant to adequately cover this gap.

Objectives

The general objective of this study is to assess the plausibility of applying the Generalized Linear Mixed Model (GLMM) to survey data in order to identify the effects of Fixed and Random (G- side and R-side) effects factors that significantly determine impact of illness/death on the dependent members of households.

The specific objectives, therefore, are to:

1. Review the theory of the Generalized Linear Mixed Model (GLMM) with the view to applying to survey data
2. Identify the factors that are associated with dependents' likelihood of re-allocating their time to care for the sick from the MMPL and RMPL models and to find which of the two models is better for this objective
3. Identify the factors that are associated with household dependents' likelihood of working harder to substitute for lost household income, as a result of illness/death, from the MMPL and RMPL models and to find which of the two models is better for this objective.
4. Identify the factors that are associated with dependents' likelihood of leaving job to care for the sick, as a result of illness/death, from the MMPL and RMPL models and to find which of the two models is better for this objective
5. Identify the factors that are associated with households' reducing expenditure, as a result of illness/death, from the MMPL and RMPL models and to find which of the two models is better for this objective.

Population under Study

This section describes the general population and the target populations covered in this study.

The Socio-economic Characteristics of the Population of the Greater Accra Region of Ghana

The survey was conducted in the Greater Accra Region of Ghana. The Greater Accra Region is located in the south-central part of the country. It shares common borders with the Central Region on the west, Volta Region on the east, Eastern Region on the north and the Gulf of Guinea on the south. It is the smallest of the 10 administrative regions of Ghana, occupying a land surface area of about 3,245 square kilometres or about 1.4 per cent of the total land area of Ghana. It happens to be one of the fastest growing urban areas within the West African sub-region. It has a coastline of approximately 225 kilometres, stretching from Kokrobite in the west to Ada in the east.

The political administration of the region is through the local government system that derives its authority from the 1992 Constitution of Ghana and the Local Government Act 1993 (Act 462). Under this administration, the region is divided into six areas/districts with their capitals, as at 2009. They are Accra Metropolitan Area (AMA), Accra; Tema Municipal Area, Tema; Ga East District; Ga West District; Dangme West District, Dodowa; and Dangme East District, Ada-Foah. Each administrative area is under the control of a Chief Executive representing central government but deriving his/her authority from an Assembly, headed by a Presiding Member elected from among the members themselves

Demographic Characteristics

The population of Greater Accra increased by almost five times between 1960 and 2000. Its share of the total population of the country steadily increased from 7.3 per cent in 1960 to 15.4 per cent in 2000. Though the male population grew by more than five times from 1960 to 2000, over the same period, the female population grew much faster, Table 1.

Table 1: Changes in key Demographic Characteristics of the greater Accra Region

	1960	2000
Population	491,817	2,905,726
Percentage share of total population of Ghana	7.3	15.4
Male population	261,547	1,436,135
Population density	151.6	895.5

Source: Ghana Statistical Service

The region has remained the most densely populated region in the country since 1960. Population density (measured as the number of persons per square kilometre) also increased by almost six times from 1960 to 2000. The region's population density, in fact, doubled between 1984 (i.e. 441) and 2000.

Economic Characteristics

Out of a population of 1,945,284 persons aged 15 years and older, 1,377,903 (or 70.83 percent) are economically active, while 567,381 (or 29.17 percent) are not. Among the economically active population, 82.6 per cent worked, 4.0 per cent had jobs but did not work and 13.4 per cent were unemployed during the seven days before the 2000 census night. It is noted that the proportion unemployed (13.4%) in the Greater Accra Region is slightly above the national figure of 10.4 percent. The proportion of employed persons in 1984, for the same region, is 92.3 per cent. In 2000, a slightly higher proportion of males (87.0%) than females (85.7%) was employed, while the reverse was the case in 1984.

Students (35.9%) and homemakers (25.8%) form the highest proportions of the non-economically active segment of the population. Persons who could not work on account of old age constitute 6.5 per cent and the retired/pensioners make up 4.9 percent. A large proportion (15.7%) of inactive population includes beggars, voluntarily unemployed and persons living on independent income or remittances, as at the 2000 Population Census. Of the economically active males aged 15 years and older, 83.5 per cent worked while 3.5 per cent had jobs but did not work. The corresponding figures for females are 81.7 per cent who worked and 4.5 per cent who had jobs but did not work. Females (13.8%) tend to be slightly more unemployed than males (13.0%). For the non-economically active population, students form the largest group. As expected, the proportion of male students (42.0%) is higher than that of females (30.7%), while females are about one-and-a-half times as likely as males to be homemakers.

Occupational Characteristics

The occupational structure shows that 42.0 per cent were engaged in sales and service occupations, with 24.7 per cent as production, transport and equipment operators. As expected, the region has a larger concentration of professional and technical workers (10.8%) compared to the national figure of 6.5 percent. On the other hand, agriculture, animal husbandry and forestry, fishermen and hunters, do not feature as prominently (9.1%) as is the case for the country as a whole (49.1%) per the 2000 population census.

There are sex differences in terms of type of occupation. The four largest male occupational groups are production, transport operators (29.6%), sales workers (19.4%), clerical and related workers (14.4%) and professional, technical and related workers (13.4%). In contrast, females are mainly sales workers (42.0%), production, transport and equipment operators, (19.5%) and service workers (13.9%).

Employment Characteristics

More than half (51.8%) of the economically active population are self-employed, while 32.6 per cent are employees. A much larger proportion of females (62.6%) than males (41.6%) are self-employed. Males were 1.5 times more likely than females to be employed.

Institutional Sectors of Employment

The private informal sector employs 62.3 per cent of economically active persons, followed by the private formal sector (23.3%) and the public sector (11.5%). Whereas 69.1 per cent of females are in the private informal sector, the corresponding figure for males is 55.8 percent.

This phenomenon is partly explained in terms of relatively low female educational attainment. A larger proportion of males than females is employed in the formal (public and private) sector.

Ethnicity

The Bureau of Ghana Languages provided the classification of ethnic groups in Ghana, which has been used since the 1960 census. Such classifications are only generic descriptions to cover a broader spectrum of ethnic groupings. In AMA, Ga and Tema there are three predominant ethnic groups namely; Akan, Ga-Dangme and Ewe.

The most predominant ethnic group is Akan, accounting for over 40 per cent of the population in AMA, Ga and Tema. While the Ga-Dangmes are the second most populous in the AMA (29%), Ewes are the second most populous in the Ga district (25.5%).

Health Characteristics

There are hospitals located within all communities in the AMA and in 14.9 per cent of communities in Tema. On the other hand, hospitals are now available in the Dangme West and Dangme East districts of the region. For instance, the maximum distance to the nearest hospital in the Dangme West District is 49 kilometres. Of more than 2000 doctors nationwide in 2003, over 50 per cent (53.9%) lived and worked in the Greater Accra region, the population of which is only 15.4 per cent of the country's total population. Greater Accra had a total of 1082 doctors, 864 of whom were in the public sector (Ministry of Health and Ghana Medical Association, 2003). This is not markedly different from the number of traditional healers (1,207) over the same period.

The population per doctor for Greater Accra was 2,686, far better than the national average of 1 doctor to 9,418 people in 2003. This is however deceptive in terms of the spread and availability in the region, because 991 (94%) of the 1082 doctors were in the Accra metropolis, with another 83 (7.7%) in the Tema municipality. Ga, Dangme East and West, between them, shared only 8 doctors at the time. Of the 991 doctors in the Accra metropolis, 483 (48.7%) work in the Korle-Bu Teaching Hospital while another 116 (11.7%) worked in the 37 Military Hospital. If one took into consideration the substantial numbers working in other major hospitals such as the Police, the Trust, the Psychiatric and the Ridge Hospitals, as well as those in private practice, this left relatively few doctors to serve the rest of the city, the metropolitan area and the region as a whole at the time.

Indeed, outside of Accra, the population-to-doctor ratio was worse than some of the most rural districts in the country. This contrasted sharply with the number of traditional healers, who were within easy reach in all the districts.

Design of the Study

This subsection provides the background information on the populations under study. It also focuses on the general impact of illness and death due to HIV/AIDS, as compared to impacts of death due to other illnesses than HIV/AIDS, within selected households of the Greater Accra Region of Ghana.

The study was conducted in the Greater Accra Region of Ghana for convenience purposes. The target population for this study was three-fold: – households that had a current or recent experience with HIV/AIDS morbidity and mortality, households with a recent morbidity and mortality due to other causes apart from HIV/AIDS, as well as households that had not experienced any recent morbidity and mortality. For households with an HIV/AIDS-related morbidity and/or mortality experience and those with morbidity and mortality due to other illnesses, the retrospective period was one year prior to the time of data collection and for households without any morbidity or mortality, the retrospective period was three months prior to the time of data collection. The three-month reference period was chosen for the third group in order to have some cases of ‘no morbidity and no mortality’ since it is quite rare to have cases of ‘no morbidity and no mortality’ in a household in a year. For the other two target groups also, the retrospective period of one year was used in order to have enough respondent households. The researcher was of the view that if a longer period was chosen, respondents might not be able to remember events that took place in their households vividly.

Outline of Thesis

This thesis is made up of six chapters. The first chapter presents a background to the study. The tail end of the first chapter highlights the objectives of the study. The review of related literature is presented in Chapter Two. In this chapter literature pertaining to the measuring of social and economic impact of HIV/AIDS and other illnesses on households has been reviewed. In addition, household coping strategies have also been looked at. Literature is also reviewed on the applications of statistical modelling, in particular Generalized Linear Mixed Modeling (GLMM), to disease impact data. Chapter Two is a review of literature on the impacts that diseases have on households.

Chapter Three presents the methodology utilized. It covers the methods used in collecting the data, the theoretical background to statistical techniques used in the study, and some methods used in managing the data. It further looks at

Generalized Linear Mixed Model building strategies in theory and with the application of PROC GLIMMIX in SAS®. Chapter Four focuses on both preliminary analyses of the data generated for this study (in SPSS®) and the GLMM (in SAS) together with the discussion of the findings.

Chapter Five then draws down the curtain by presenting the summary, conclusion and recommendations of the study.

CHAPTER TWO

REVIEW OF RELATED LITERATURE

Introduction

This chapter explores some of the studies on the assessment of the social and economic impact of poverty and HIV/AIDS on households. Central to understanding the social and economic impact of poverty and HIV is to understand what goes on in the household affected by the disease. This chapter tries to understand the dynamics that arise when the household is affected by HIV. In this study, a household is considered an HIV when a member is infected with HIV. This is used in order to identify households with HIV-positive cases. Moreover, any impact due to the disease is borne by all household members, directly or indirectly. This chapter reviews literature on the economic impact of HIV/AIDS on households and how this impact is measured.

Measuring Social and Economic Impact of HIV/AIDS on Households

Literature abounds on the social and economic impacts of HIV/AIDS on societies and economies. However literature on statistical modelling of the social, economic and demographic impact of the epidemic on households is comparatively small and limited (Barnett et al., 2001). Barnett et al. (2001), who were major contributors at the maiden Global Conference on AIDS, further stated that very limited effort has been made towards understanding the impact of HIV/AIDS at the household level. To this end, studies towards the statistical modelling of the social and economic impact of the epidemic on households is very relevant especially for designing interventions to minimise or curb its direct and indirect impact on households.

HIV/AIDS is reaching a stage at which its related morbidity and mortality are increasing at very rapid rates (Dorrington et al., 2001). In countries which have been hardest hit, typically African countries, adults whose responsibility it is to fend for the non-working members of the household (i.e. the young and elderly), are sick and some bedridden (Gross, 1997). This has brought much pressure on the young and elderly, who should have been cared for by the sick, now having to care for these sick breadwinners. The situation can exert untenable pressure on households in their struggle for survival. Poor households are often the worst hit and more vulnerable to the long-term effects of HIV/AIDS (VanLandingham et al., 2000).

VanLandingham et al. (2000) also identified other characteristics, which aid coping with HIV morbidity/mortality such as those of the community, including attitudes towards helping needy households and the general availability of resources. In addition, they also indicated that because of the protracted nature of the disease, the impact of an AIDS death may result in a lengthy depletion of household resources thus resulting in greater and more enduring hardship than some other causes of death. According to them, there is some evidence that women bear a heavy burden of the household impact at all stages from early childhood, when they may be less well-nourished or removed from school to save money for care costs of a sick parent, through to stigmatization on the death of a husband, and finally suffering lonely and impoverished widowhood. However, these three afore-mentioned studies considered largely economic impacts of HIV/AIDS.

Some studies on HIV/AIDS hold the view that the menace impacts households at two main levels. These are the economic and social levels. For instance according to VanLandingham et al. (2000), on the economic level, households and the surviving members have to pay for medical costs and funeral expenses and, if the deceased was a breadwinner, there will be further financial impacts in the form of loss of income. They also indicate that at the social level, households have to deal with issues around stigmatisation, social exclusion and disintegration of family structures and social support networks. Women, especially, are overburdened with care and support roles.

HIV/AIDS mortality can change the demographic structure of the household, reverse the roles of the members, exacerbate poverty, rob children of their parents thereby creating more orphans (Wijngaarden and Shaeffer, 2005) infringe on the basic rights of the child in areas such as education, food, nutrition, health and other social benefits. Unless households are strengthened and empowered through focused interventions, poor households are likely to fall deeper into poverty for the generations to come.

HIV/AIDS continues to exceed all expectations in the severity and scale of its impact on households and countries in general. Piot et al. (2001) predicted over a decade ago that AIDS constitutes one of the most serious crises currently facing human development, and threatens to reverse progress in the mostly affected countries by decades. There is no reason to believe that Africa as a continent is currently not feeling the effects of this pandemic especially considering the death toll and ever increasing number of orphans produced as a result of HIV/AIDS.

HIV/AIDS significantly impacts households and their ability to cope with the epidemic. According to Piot et al. (2001), household impact is one of the points at which AIDS demonstrates its effect. They assert that the disease exacerbates and prolongs poverty in every context. For example in poorer households, AIDS takes a greater proportion of expenditure, and limits access to food and health care. In education, it has a negative impact both on the supply of teachers and on the capacity of children to continue in school (Wijngaarden and Shaeffer, 2005).

Poverty and HIV/AIDS

About 63% of global AIDS cases occur in Africa. Thus, the menace is affecting sub- Saharan Africa more severely than any other parts of the world (Lugalla et al., 1999). Poverty and HIV/AIDS are obviously cause for concern for the African continent. Sub-Saharan Africa is the only region of the world where the proportion of people living in extreme poverty is increasing. According to Jooma (2005), the number of Africans living below the poverty line (less than 1 US dollar per day) has almost doubled from 164 million in 1981 to 314 million people today. Jooma (2005) further states that 32 out of 47 African countries are among the world's 48 poorest nations. The impact of extreme poverty is felt even more at the household level, according to him. Households may find themselves spiralling into extreme poverty, making it impossible for them to assume their "normal" functioning.

In trying to trace the impact of HIV/AIDS to the early beginnings of the discovery of the menace, Lwihula (1992), in Lugalla et al. (1999), linked the AIDS epidemic with the years of economic crisis in the early 1980s that saw the scarcity of essential commodities. According to him, these economic hardships intensified poverty, destabilized families, and increased people's movements between countries. The situation widened the web of sex networking, and in this way facilitated the early rapid spread of HIV. This is further supported by Lugalla et al., (1999) when they said that HIV/AIDS does not occur in a vacuum but rather in a social context.

Understanding poverty within the context of HIV/AIDS is critical as it is viewed in this chapter as both a risk factor for HIV infection and the consequence of it. Cohen and Reid, (1998) say that as a risk factor, poverty is associated with weak endowments of human and financial resources such as low levels of education, low levels of literacy and few marketable skills, generally poor health status and low labour productivity. Cohen and Reid (1998) go on to state that the inability to attract endowments, through engaging in income generating activities by adults, as a result of HIV infection, morbidity and mortality sinks poor households into even deeper poverty. Poor households may find it even more difficult to exonerate themselves from dire poverty for many more years and generations to come. Poverty, as a consequence of HIV infection could see the poor adopting various mitigation strategies to cope with the disease that often exposes them to HIV infections. Cohen and Reid (1998) argue that it is not surprising that the poor adopt behaviours that expose them to HIV infection.

Whiteside (2002) suggests that illness and poverty affect household resources and income. In the face of the rising effect of HIV/AIDS, one sees rising costs of medical care/treatment, and an increased need for nutritious foods. With the progression of the illness, the demand for care also rises. Children are often withdrawn from schools to care for sick adults, further compromising their basic right to education (Wijngaarden and Shaeffer, 2005). The deprivation of education could place the household at further long-term risks for poverty, lack of skills and disempowerment. The latter result is a cycle of household impoverishment that may take decades to reverse.

Studies have shown that lack and/or limited education and skills also appear to influence vulnerability to HIV infection. Shisana and Simbayi (2002) in a national household survey in South Africa, found that those with tertiary educational qualifications had lower rates of HIV infection than those with only Primary and Secondary school level qualifications. The assumption here might be that people with the necessary educational qualifications, thus acquiring economic independence or freedoms for survival do not engage in risky behaviours more than those with limited education.

Cohen and Reid (1998) also argue that HIV intensifies poverty, leads to its persistence and over time generates a culture of poverty. When parents are sick and die from AIDS-related complications, little or no transfer of skills and knowledge to the younger generation takes place. The circle of poverty is likely to repeat itself and felt over generations, according to him. Barnett et al. (2001) argue that interventions to mitigate the effects of the pandemic on the rising generations are needed. Loewenson and Chikumbirike (2005) also hold the view that persistent poverty leads to what is termed "new variant famine" where chronic poverty and ill health are increased, thereby reducing household mechanisms and resources for coping with illness and mortality and further undermining long term prospects for food security and household well-being.

The sexual activities that men and young women engage in, in exchange for money, have been described by Nicoli Nattrass (2004) as "sexual economy". The participation in the sexual economy activities, as a result of poverty, places young women, in particular, at higher risk of HIV infection transmission. Akeroyd (1997) asserts that sexual culture places women in a vulnerable situation regarding HIV infection. He goes on to say that poverty exacerbates it by encouraging women to engage in sex as an economic strategy for survival. Dixon-Fyle and Mulanga (2004) support Akeroyd's (1997) view by stating that young women sell their bodies to help families, and men take advantage of the opportunity, or express feelings of powerlessness and despair through sexual violence.

Lugalla et al. (1999) report that gender inequality and poverty prevent women from exercising their ability to fulfil their socially designated responsibilities, and therefore debases them, often forcing them into prostitution. Shelton et al. (2005) also commented that the poor, especially women, are vulnerable to sexual exploitation because HIV prevalence is partly a function of survival. They further contend that people with HIV eventually tend to lose wealth because of loss of employment and increased expenses related to the disease, thus blunting a positive relation between wealth and HIV.

Studies conducted by Lugalla et al. (1999) and Munyako (1994) indicate that a decline in government expenditure on health in many African countries translates into an increase in a number of untreated STDs that are known to facilitate the rapid transmission of HIV. This could have serious long-term health implications resulting from the rapid spread of HIV.

According to Verner and Alda (2004) Children raised in poor households face a large risk of achieving a low level of educational attainment and educational attrition. Girls especially are removed from school as a coping strategy, and also because the girls education is viewed as “less of a priority”, since it is expected that they will marry and belong to another family (Grant and Palmiere, 2003). This is also largely due to economic factors such as loss of income due to HIV/AIDS amidst high education costs and the direct costs like school fees, textbooks and uniforms.

From literature discussed so far, it is evident that HIV/AIDS appears to be associated strongly with poverty and has increased the depth of vulnerability of those households already vulnerable to shocks. HIV/AIDS has acted to intensify the disadvantages imposed on the poor households and communities. However HIV/AIDS is said to have both direct and indirect impact on its affected households. These are largely in the form of costs incurred by households as its members progress from HIV infection, through AIDS-related illnesses and ultimately death. The economic impact of HIV/AIDS morbidity and mortality on households are therefore commonly analysed in terms of direct and indirect costs

(Bollinger and Stover, 1999) and are usually reported as proportions of income and expenditure

Direct Economic Impacts

In HIV/AIDS studies, direct costs represent actual expenditures on treatment (Booysen et al., 2002; Danziger, 1994; Bowie, 1996) and funerals (Pitayanon et al., 1998; Booysen et al., 2002). While HIV/AIDS crosses all socio-economic groups, its economic impacts are greater on the poor, powerless and marginalized (Grant and Palmiere, 2003). This stems from the fact that from the time of diagnosis, poor households feel the economic impact of the disease. Wyss et al. (2004) found in their study in Chad that the average costs of AIDS to patients and their families are very high. On the average, a household spends the equivalence of USD78.6 a month directly on AIDS treatment and care. Cross (2001) in her study on rural households in South Africa further asserts that the de facto per capita income may fall to as low as R50 per month. The households therefore spend considerable amounts of money on consultation and treatment fees, and transport. Households see a greater spending on health care and associated costs (Save the Children, 2004; Wyss et al., 2004).

The chronically ill person is often unable to work leading to reduced income and output in agricultural production. Chronic illness coupled with the need to care for the ill, by other household members, takes valuable time away from productive activities leading to double loss of income thus exposing households to risks such as food insecurity and exposure to HIV transmission (Save the Children, 2004). In addition, De Waal and Whiteside (2003) have found that diversion of labour coupled with the care of children orphaned as a result of the death of their parents to AIDS related diseases further impoverishes the household.

HIV/AIDS strikes persons at the prime of their lives thus exerting a heavy toll on the economic well-being of the household. The death of a productive member comes with a reduction or loss of income (Cross, 2001); Save the Children,2004) and absence of savings and other assets to cushion the impact of illness and death (Cohen, 1993). For households that are solely dependent on agriculture, the death of the member means that the contribution to agricultural production and income from that person is permanently lost. However, this may also be the case for people working in the industry.

Grant and Palmiere (2003) found in their study in Bulawayo (Zimbabwe) that HIV/AIDS affected households experience a 40% drop in household income, which is bound to impact the decisions and the psychological outlook of the household. The lack of time is viewed as a contributory factor to the dip in household income. Although households attempt to diversify, they are unable to add a lucrative income-generating project. Households may be forced to change their livelihood strategies to counter the impacts of the loss and reduced household income. As was found in Grant and Palmiere (2003), households were forced to cut back on their livelihoods to accommodate a lower average monthly income, and an increase in the number of people living within the household. This effectively means that households sink deeper into poverty and likely chances to avert the economic impact are very low or non-existent for some very poor households.

The HIV/AIDS epidemic undercuts the ability of the households to cope with shocks. Assets are likely to be liquidated to pay for the costs of care. Sickness and caring for the sick prevents people from migrating to find additional work (Wiggins, 2005).

Indirect Economic Impacts

According to Booysen et al., (2002) and Cohen and Trussell, (1996), indirect costs are commonly associated with loss of earnings to the sick person, the deceased and/or the caregiver. In addition, People Living with HIV/AIDS (PLWHA) may suffer from considerable stigmatization in their homes, communities and workplaces when their HIV+ status is known. This may lead to various forms of social and political discrimination/exclusion including reduced chances for employment, in some cases dismissal from work, and insensitive and biased institutional policies. Lau and Wong (2001) have found that almost 20% of companies in their study would dismiss HIV+ employees to avoid anxiety and unrest among the rest of the staff. They further found that HIV+ employees would be transferred to other posts/positions against their will once their HIV+ status is known. This indicates that stigmatization may impact the financial resources of the household that could otherwise be generated through formal employment.

Grant and Palmiere (2003) have it that following the gender bias argument, women come out worst in terms of income generating activities available to them. According to them, because there is a general expectation on women to care for others including the sick, valuable production time is lost thereby impacting on the economic ability of the household to offset the ill effects of the pandemic. Wyss et al. (2004) found that time lost due to illness was 15.8 days a month and family members spend time caring for the ill person instead of engaging in income generating activities. Household members provided assistance at an average of 8.3 days thus abandoning their daily activities or occupations. Average monthly productivity loss attributable to AIDS equalled 21.6 days per household.

The HIV/AIDS epidemic reduces farm production and incomes. In farming, labour is lost to sickness and death, as well as to the time taken by those caring for the sick. Affected households plant smaller areas and use less intensive production methods (Wiggins, 2005). Capital to buy inputs is likely to be spent first on medicines, visits to hospitals and eventually on funerals.

Household livelihood is a critical factor in understanding the impact of poverty and HIV on the overall functioning of the household and its ability to provide for the basic needs of its members. The concept of livelihood is therefore multi-facetted in that it considers the activities that the household engages in and the outcomes thereof. It also reveals the interconnectedness and/or the interplay between the household activities, environmental and social institutions in the community/society that determine the outcome or livelihood of the household.

Effects of HIV/AIDS on Household Food security

Food security is an important element for the survival of any household across the spectrum of wealth. Households affected by HIV may find it difficult to maintain their food security. HIV exerts tremendous pressure on the household's ability to provide for basic needs like food.

Agricultural activities contribute to the welfare of households in two ways. Firstly, the production of food crops and ownership of livestock contributes to food security and secondly it provides income (Samatebele, 2005). HIV/AIDS has a retrogressive effect on the agricultural sector of poor countries, and therefore on households since agriculture in these countries are basically household-based. It therefore becomes difficult differentiating between household income and expenditure and those of agriculture (Topouzis, 2003). A major impact on agriculture includes the depletion of human capital, diversion of resources from agriculture, loss of farm and non-farm income together with other forms of psychological impacts that affect productivity (Jooma, 2005). De Waal and Whiteside (2003) further assert that households with a chronically ill person see an income reduction of between 30% and 35%.

Food shortages in Southern Africa are ongoing problems, and long-term projections suggest that regional food production per capita is likely to diminish into the future (Rosegrant et al., 2001). Food crisis is undoubtedly made worse and malignant by a fully-fledged HIV/AIDS epidemic. The disease leads to competition within a household for its resources – money and productive capacity must compete between care-giving and health-care costs on the one hand and agricultural inputs and labour on the other (Stewart, 2003). Food shortages could severely hamper the health of HIV infected individuals. The quality of life of people infected with HIV has implications for national productive capacity, for the stability of family and social structures (Stewart, 2003).

Poor nutrition is often linked with adverse outcomes in HIV/AIDS. Poor nutritional status is linked to vulnerability to progression from HIV infection to mortality (Bates et al., 2004). Poor nutrition weakens the body's defence against infection and infection in turn weakens the efficiency of absorption of nutrients. Micronutrient deficiencies undermine the body's natural defences against infections, thus contributing further to the vulnerability to HIV infection (Nattrass, 2004). Households experiencing food shortages as a result of poverty and effects of HIV/AIDS increase the chances of fast progression of the illness and inevitable death of the ill person.

Given that HIV/AIDS leads to poverty and malnutrition, there is thus a good reason to assume that poverty helped hasten the spread of HIV in sub-Saharan Africa (Nattrass, 2004). Parasite infection, mainly malaria and intestinal parasites undermine the nutritional status and compromise the immune system yet further, effectively exhausting it. Such parasite infections are endemic in Africa but the situation is made worse by inadequate health care and infrastructure. It must be noted that inadequate health care and infrastructure are a function of poverty and low levels of development and these leave most parasite infections untreated (Nattrass, 2004).

HIV/AIDS and Household Health

HIV/AIDS is having a devastating effect on health in many countries in sub-Saharan Africa. (Zabaa et al., 2004) The report "World Health Organization's (WHO) Commission on Macroeconomics and Health" (2002) sees ill health as a dimension of poverty, and advocates investing in health as a means of working towards poverty reduction and raising living standards of the poor (Bloom and Canning, 2003). Bloom and Canning (2003) further contend that the physical body is the poor people's main asset, but one with no insurance and ill-health therefore imposes a higher level of risk on the poor when the principal asset is struck down by a disease. They cannot earn the money needed to provide themselves (and usually others too) with food or medicine, and the health shock is likely to be catastrophic.

Increased adult morbidity and mortality associated with HIV infection are likely to have important consequences for households, communities and health systems (Ngalula et al., 2002). One such consequence is economical, as households have to pay for health care services. A study in Tanzania revealed that terminal illness associated with HIV/AIDS is associated with high levels of modern and traditional levels of health services use, mainly because of the longer duration of the illness (Ngalula et al., 2002). The more an HIV infected person is suffering from morbid acuteness of the disease, the more likely it is that the sick person will seek help from health institutions. In some Central African states, 60% of hospital beds are occupied by patients with HIV/AIDS related conditions (Sibanda et al., 2003).

Child Rights

HIV/AIDS predisposes children to violation of their basic rights. Children are dependent on adult members of the household for food security. Failure of the households to provide children with nutritious foods may hamper their nutritional status thus placing children at risk of various infections that would undermine their health status.

Chronic illness in children can lead to physical, social and developmental delays (Warwick et al., 1998). This can contribute to longer-term challenges that have to be addressed by households and family members. If the chronic illnesses remain untreated, it could lead to further impoverishment of the household in the long-term. Children may be ill and unable to go to school and attain better educational qualifications that could be utilized to the betterment of the quality of life of the household.

In fact, existing studies show that children raised in poor households (most of which come about due to HIV/AIDS) face a large risk of achieving a low level of educational attainment and dropping out of school (Verner and Alda, 2004). The intergenerational transfer of low levels of education is high in households hardest-hit by HIV/AIDS.

Where the impact of AIDS has been greatest, and where there are few, if any, adults to care for the bereaved children, a few households may be constituted of children alone (Warwick et al., 1998). Children are therefore deprived of warm and caring homes and forced into situations where children have to lead households irrespective of their experience and need to be cared for as children. Warwick et al. (1998) further assert that poverty may force childcare to be provided outside preferred social networks.

HIV/AIDS is changing the age distribution of the labour workforce with an increasing number of children facing economic uncertainty and hardship. The early entry of orphans into the labour workforce exacerbates the worst form of child labour, and the epidemic is forcing older persons back into the workforce due to economic need (Dixon-Fyle and Mulanga, 2004).

Coping Strategies

Households respond in various ways when trying to cope with or mitigate the effects of HIV/AIDS. Various authors have written on the coping strategies of households some of which are presented in this section.

Household livelihood diversification is defined as a process by which households construct an increasingly diverse portfolio of activities and assets in order to survive and to improve their standard of living (Ellis, 2000). Diversification is generally recognized as an important strategy for decreasing livelihood vulnerability. The poor are left with little chance for survival hence diversification gives them an opportunity for revival and/or recovery (Niehof, 2005; Whiteside, 2002).

As a result of desperation for household survival following the severe socio-economic impacts of HIV, households may sell their moveable assets to pay for medical costs and funeral expenses. In time, households delve deeper into poverty and impoverishment as a result of the sale of their assets and may reach a point at which economic recovery becomes impossible. For example, in agricultural communities, once households have sold all their livestock, they may resort to selling their tools which can mean that they are even unable to sell their labour, since they do not have the implements with which to work (Grant and Palmiere, 2003; Whiteside, 2002; Cross, 2001).

Households revert to borrowing credit from the informal sector to offset the immediate impact of HIV and poverty. This offers a short-term solution to long-term problems households are faced with. Households need to be assisted to engage in sustainable activities to deal with the long-term effects of HIV and poverty (Cross, 2001; Grant and Palmiere, 2003).

Characteristics of HIV Affected Households

Save the Children (2004), a leading UK charity working to create a better future for children and young people, identified three sets of circumstances to define affected households viz. chronic illness, death and support of orphans.

Chronic Illness

In a chronic situation, the HIV infected person is unable to work and this inability contributes to reduced household income. Shortage of labour as a result of illness could lead to role restructuring within the household. Drimie (2003) notes that women, the elderly and young people often assume greater burden of ensuring household survival, in addition to taking the burden of caring for the ill.

Caring for the ill could mean time taken away from productive activities, land utilization and education. The above factors create a cycle of dependency among members of the household. The severity and amount of strain put on other members to care for the ill and the members' endurance will determine the duration of survival of the household. The inability to endure such pressure may render the social support networks inoperable. Children may find themselves in the centre of this situation when they are withdrawn from school to fill the gaps. The situation would then severely compromise the right of the child to education.

When children are withdrawn from school to care for the ill and fill the gaps where additional labour is needed, the household may be faced with “double loss of income”. The ideal of utilizing education as a means of fighting off poverty, in the longer term, is then diminished. The household is likely to go deeper into poverty with little or no hope of recovery.

Death

In the event of the death of the infected household member, who may be a breadwinner, the contribution to agricultural production or household business and income from that member is permanently lost. Studies have shown that households’ land cultivation has reduced as a result of the death of the breadwinner and sometimes adults who were actively involved in agricultural production (Drimie, 2003). De Waal and Whiteside (2003) are of the view that AIDS puts households at increased vulnerability to famine.

The direct costs of death due to AIDS are substantial. Firstly, the household would have used substantial amounts of money in health-related costs prior to the death of an ill person. By the time a person dies the financial resources of the household might have been exhausted already. The immediate economic impact in the event of death to AIDS-related complications on the household is the funeral expenses.

Grant and Palmiere (2003) argue that the primary economic cost of HIV/AIDS-related death is the foregone income of the deceased. This is assuming that the deceased person was economically active and contributing to the livelihood of the household.

Support of Orphans

Households across the entire spectrum of wealth can take in orphans reflecting the facts that HIV/AIDS affects all types of households. Poor households are the hardest hit as they are forced to make ends meet with the little resources they have. The addition of orphans into an already impoverished household drains the household financial resources. However, taking in an orphan, depending on his/her age, gender and health, may bring a net economic benefit to household income or food production (Save the Children, 2004). So taking in orphans is not necessarily a bad thing since it can enhance the livelihood of the household. However, in many African societies, tradition demands that households take in orphans of relatives regardless of whether they have the means to support them.

Loss of one or both parents, depending on specific cultural traditions and levels of household endowments, is likely to decrease physical, emotional and mental welfare of the child (Barnett et al., 2001). In poor households where food consumption is reduced for economic reasons, this may severely impact the physical and health status of the child. Some children may have not been immunized because parents were sick and unable to access health services for their children.

The inclusion of orphans into an impoverished household has an impact on the household food security (Save the Children, 2004). Younger children require more care and support than older children. There is a need for interventions to mitigate the effects of HIV/AIDS and poverty on orphans and households.

Effects of HIV/AIDS on Household Structure and Relations

As individual productivity is affected, costs diverted to the illness are likely to have an impact on household food security, where the husband is falling ill and his wife takes over a caring role, while being less involved in her traditional 'productive' activities (Cohen, 1993). Children become important contributors in both care and production, which results in reduced schooling for children, especially as financial resources become scarce. According to TASO personnel (TASO is an AIDS counselling service in Tororo District, Uganda), only one in five children from HIV-infected households stay in school (Topouzis and Hemrich, 1994). Furthermore, health care for the rest of the family may suffer, as it becomes unaffordable (Cohen, 1993). Barnett and Blaikie (1992) illustrated how over a period of ten years, a household reacts to the course of AIDS, which affects different members of the family and eventually leaves the children as orphans, who have to labour for other households, as their own land has been abandoned.

Widows and widowers

A UNDP analysis in three districts of Uganda, Kabarole, Gulu and Tororo, revealed that that far more women had lost their husbands than vice versa. A man tends to be relatively cushioned after his wife's death, as he can either rely on his other wives in the case of polygamy, or may even start looking for a new wife while the first one is ill (Topouzis, and Hemrich, 1994). For women, widowhood leaves them highly vulnerable to poverty.

According to a UNDP study, a widow's future perspectives depend greatly on the way her husband's family perceives his death. It is not uncommon that his wife is blamed for his illness and is accused of promiscuity and immorality. If she then feels forced to leave the area, she may become a migrant, with associated risks of poverty and insecurity (Topouzis, and Hemrich, 1994). In this situation, a widow is likely to seek a new partnership to secure her livelihood, but this puts her future partner at risk, if she is infected and is frightened to let him know about her past.

The experience of TASO, illustrates the importance of working with both men and women according to their specific needs, for example in the case of wife-inheritance. For example, counsellors and widows reported that it was not uncommon that brothers-in-law would disregard the dangers of infection and would even abandon HIV-infected sisters in law who refused sex on the ground of not wanting to infect the extended family (Topouzis and Hemrich, 1994). An infected husband can help prevent poverty for his family, if he agrees to write a will, which grants their common property to his wife. In order to challenge social traditions such as wife inheritance the support of male members of the community is required. They have to be mobilised at the same time as giving women support and economic alternatives to enable them to resist the practice.

Orphans and Child-headed Households

The understanding that women are heavily infected and will eventually fall ill has created an additional problem, that of a rising number of orphans. UNAIDS (1998c) estimates that 1.1 million orphans live in Uganda. A survey of

1797 rural and urban households in six districts in the South of Uganda by Ntozi et al. (1997) showed an overall orphanhood prevalence of 42.7 percent, reaching as high as 64 percent in Masaka District. A general national census in 1991 based on a random sample revealed an overall average prevalence of orphans nationwide of 10.7 percent. This does not clarify the extent to which orphanhood is attributable to HIV/AIDS per se.

However, according to the Ntozi et al.'s, (1997) study, in 54 percent of cases studied the death of a parent was AIDS related. In Masaka District, this rose to 82 percent. The data also revealed that more children had lost their fathers rather than their mothers with an overall sex ratio of 159 fathers to 100 mothers. This was the case for AIDS and AIDS-related diseases (male-female sex ratio of 1.2 and 1.5, respectively) but more so for other causes of death. The fact that more men than women had died from AIDS confirms that more men are dying and that within families; it is usually the man who presents first with the disease. AIDS-related paternal orphanhood (i.e. children who have lost their fathers) accounted for almost 40 percent of all causes of death of parents. There is no indication as to how many children had lost both parents, which is likely to be the case once the mother has died. However, the author discusses 'surviving fathers', which indicates that at least some wives acquired the disease (and died) before their husbands.

The gender implications of the analysis of the demographic impact of HIV/AIDS lies in the sudden recognition of reproductive tasks of women, as they disappear as carers, especially when the mother has died. The UNAIDS (1999) definition does not even consider paternal orphans as orphans, with the implicit assumption that mothers will just get on with the care. Especially in the case of AIDS, there is a strong likelihood that the surviving mother will be infected and therefore require assistance to secure the livelihood of herself and her children. Surviving fathers are more likely to re-marry (Topouzis, and Hemrich, 1994) or give their children up into the care of grandparents (Ntozi et al., 1997).

When the extended family takes over part or all the care of the children, this raises the question of who will do the bulk of the work. It seems more than likely that the main carers are women, who are then consequently restricted in their own activities to secure their own livelihood (Topouzis and Hemrich, 1994; Grundfest-Schoepf, 1991). Another problem is the ages of alternative carers who may be either very young or very old (since AIDS deaths peak in early adulthood) and are themselves highly vulnerable to poverty (Sengendo and Nambi, 1997). In the Ntozi et al.'s (1997) survey sample, older siblings reported to be primary carers for the orphans in 7.2 percent of cases, which indicates a growing phenomenon of 'child-headed households'.

According to the study by Ntozi et al. (1997), paternal orphans suffer more from lack of parental care whereas maternal orphans especially lack money. Both have serious implications on the future development of the child. In general, the most striking effect for children seems to be poverty due to the processes described above, whereby children of female-headed households are most affected where widows have lost access and control of productive family resources. It is likely that girls, more so than boys, will be taken out of school, as they can easily substitute for mothers in domestic tasks.

So far, little has been done to address psychological effects for children who lose their parents to AIDS (Sengendo and Nambi, 1997). A recent study showed that the trauma of experiencing the death of one or two parents, and the lack of emotional support was felt especially by maternal orphans. They were more predisposed to physical and psychological risks and tended to be more 'externally orientated' (ibid.), which may have implications for their sexual behaviour during adolescence. The study, which was undertaken at a school in Rakai District concluded that all orphans felt less optimistic about the future compared to non-orphans, decreasing their potential to cope with their life ahead. It is estimated that 50 percent of all new infections occur among the population between 15 and 24 years and a further ten percent in children less than 15 years (Lyons, 1998), so that these children are passing directly into a high-risk phase for infection.

A final effect is on family values and traditional norms and customs, which may influence children differently according to their gender. When families are breaking up, children miss out on family-based education and guidance, especially if they are expected to mature fast and take on responsibilities (Topouzis and Hemrich 1994). It has been suggested that this leads to early sexual activity, with all its inherent dangers (ibid.). The fact that women suffer the brunt of the impact of HIV/AIDS may act as a deterrent to promiscuity, among girls, however. There is also a possibility that boys might develop a better understanding and sense of responsibility for reproductive tasks, but this can only be maintained if society is supportive. In Tororo district, increasing numbers of children run away from home to escape poverty and the stigma of being AIDS-orphans (ibid). The growing number of street-children indicates that they do not experience a supportive environment for development from their community.

Demographic Impact

The demographic impact of HIV/AIDS becomes evident if changes in the population pyramid occur which would not happen in the absence of AIDS. In order to compare scenarios, the US Bureau of the Census has produced estimates of a range of demographic indicators for 28 countries (of which all but seven are African), with or without AIDS (Advance tables of the World Population Profile, cited by Sida 1998b). According to those tables, which give information about population growth, life expectancy and crude death rates as well as child and infant mortality, HIV/AIDS has had a significant demographic impact in Uganda.

A comparison of demographic impact between Brazil and Uganda illustrates the severity of the situation in Uganda, where infant and child mortality rates are exceptionally high even without HIV/AIDS. A limitation of the model on which the estimations were based is that it does not account for HIV transmission other than through heterosexual contact, whereas in Brazil, same sex relations and drug use are major contributors to HIV transmission. This may mean that the estimates for Brazil are biased (U.S. Bureau of the Census 1998).

Fertility

Changes in fertility due to HIV/AIDS have so far not been systematically incorporated into demographic projections, although they have been discussed in literature (Zaba and Gregson, 1998; Ntozi et al., 1997).

In their evaluation of survey material from Masaka and Rakai Districts and other localities in Uganda, as well as Tanzania and Zambia, Zaba and Gregson (1998) observe that a ten-percent prevalence of HIV has the impact of a four- percent decrease in the total population fertility. Ntozi et al. (1997) came to the conclusion that fertility rates declined from 7.3 to 6.0 between 1992 and 1995 in six Ugandan districts surveyed, whereby women in AIDS-affected households showed significantly lower fertility compared to those in non-affected households. Zaba and Gregson (1998) point out that fertility estimates do not account for changes in behaviour in the general population, as knowledge about AIDS and fear of infection change their sexual behaviour.

Infant and Child Mortality

A recent evaluation of the impact of HIV/AIDS on infant and child mortality in Uganda by Ntozi and Nakanaabi (1997) found a positive association of mortality rates with parents who are educated, polygamous, formerly employed and in business. This reflects the understanding that parents with higher income are more at risk of contracting HIV/AIDS and therefore mothers are more likely to transmit the infection to children during pregnancy and breast-feeding. However, this analysis overlooks infant and child mortality not directly caused by AIDS, but rather associated with the reduction in available health care, due to increased demands on services due to AIDS (as discussed above), or because diseased parents or impoverished relatives have reduced childcare capabilities.

The Impacts of HIV/AIDS on Orphans and the Elderly

The rapidly growing number of AIDS orphans now commands the attention of a large number of researchers concerned with their care. According to the UN, the disease has resulted in more than 14 million AIDS orphans since the epidemic began (Hagen, 2002) and this number was projected to increase to some 40 million in Africa by the year 2010 (Foster and Williamson, 2000). However, these numbers are thought to underestimate the problem because they are based on a restricted definition of orphanhood. The United Nations and other major observers typically define an orphan as a child under 15 years old whose mother has died of AIDS or any other cause, thus excluding children from 15 to 17 years old and those who have lost a father to AIDS (Case, 2003). But many community programs aimed at helping children in difficult circumstances as well as those promoting rights of children often define orphans as those less than 18 years old who have lost one or both of their parents. Limited definitions of orphans hide specific problems, such as the particular needs of young adolescents and the differences between losing a mother, a father, or both (Hagen, 2002).

Thus, major questions remain about how large the orphan crisis is and how fast it is growing. Researchers who attempted to answer these questions in the early 1990s tended to use theoretical mathematical models, since there was insufficient census or morbidity data in most African countries (Gregson et al., 1994). More recent studies have been able to utilize demographic and health surveys and or other sources of empirical evidence to more closely track orphan prevalence rates and trends in various countries, although they cannot determine the cause of death of parents.

On the effects of orphanhood on children, most researches address questions of the impact of losing one or both parents on the child in terms of their educational, nutritional, health and emotional status (Zaaba et. al., 2004; and Monash and Boerma, 2004). Recent research has focused on the caregivers and surviving household heads, who are often female and elderly grandparents, and sometimes even siblings who are children themselves (Case, 2003).

In general, children who have lost one or both parents to AIDS are at risk of leaving school or falling behind their age group in school. The main concern is that families pull children out of school when the financial burden increases due to HIV/AIDS. Additionally, even before the parent dies, the child is needed more in the household to help with domestic work (Monash and Boerma, 2004). They continued by indicating that after becoming an orphan, some children stay home to take care of their siblings, and therefore do not go to school. Some studies have found that when a mother dies, younger children are less likely to go to primary school, and when a father dies, children in upper grades (where school fees are high) are less likely to go to school (Pridmore, 2008).

The long-term effect is a loss of productive human capital (Subbarao et al., 2001). However, other studies have found little disadvantage in educational opportunities for orphans (Bicego et al., 2003). Uganda provides one model for ameliorating the problem by making it possible for all children to go to primary school without fees through the Universal Primary Education Program implemented in 1997. The effect is that educational opportunities for orphans were found to be same as for non-orphans (Deininger and Subbarao, 2003). In addition, the program seems to have helped all children go to school, whether they are orphans due to AIDS or not.

Another concern is that children's nutrition may decline if they are orphans due to AIDS, although empirical studies have found conflicting results (Balyamujura et al., 2000; FAO HIV/AIDS Programme, 2002). Another study found that adding a foster child to a household had the effect of reducing per capita consumption as well as investment of household resources, which in turn negatively affected nutrition and medical services (Deininger et al., 2003).

Another complicating factor is that by the time a child has become an orphan of one or both parents due to AIDS, he or she has lived through the illness of this parent, and this has its own effects upon the well-being of the child as well as the economic situation of the family (Pridmore, 2008). AIDS can have a greater impact on children than other diseases as the surviving parent is likely to die too if also infected, and because of the enormous economic burden due to a lengthy period of illness (Crampin et al., 2003).

Many children live with their grandmothers or in child-headed households, taken care of by older siblings (Ansell and Young, 2004). In Zimbabwe, studies have shown that care giving is increasingly provided by grandparents, with an average age of 62, while a small minority of households are headed by siblings who are children themselves (Foster, 1998). Matshalaga (2002) examines the impact of orphans on the extended family system, showing that in Zimbabwe grandmothers who had traditionally “retired” from active life were drawn back into family and community dynamics through their new child-rearing responsibilities.

Although the extended family has been the focus of care for orphans and has usually been able to adequately absorb orphans within communities, especially in rural areas where extended families are more intact (Walraven et al., 1996; Kamali et al., 2010), there are signs that the extended family system is being stretched as the number of AIDS orphans rises (Preble, 1990; Danziger 1994; Nyambedha et al., 2003). One qualitative study in western Kenya found that the traditional patterns of fostering of orphaned children are not adequate for the care of the increasing numbers of children orphaned by HIV/AIDS. More and more grandparents, mostly grandmothers, are caring for numerous grandchildren, even though their own incomes are not high (Nyambedha et al., 2003).

The effects of care-giving can vary across households, but most will face a drop in living standards due to costly health care, loss of income as the sick and their caregivers drop out of the workforce, and funeral expenses, all of which can lead to debt and poverty (Danziger, 1994; WHO, 2002). Indeed, a study of household expenditure due to AIDS in Tanzania (Ngalula et al., 2002) found that the cost of medical care and funerals exceeded the annual income of many households, largely due to the long duration of this disease. In the future, the price of antiretroviral treatments determined by pharmaceutical companies and government subsidies will also play a significant role in the future impact on families (Knodel et al., 2001).

AIDS and HIV also have specific effects on the welfare of the elderly. Studies conducted in Thailand, (Knodel et al., 2001; Wachter et al., 2002; and Ainsworth et al., 2005) found that in effect, AIDS victims return to their parental homes at late stages of the illness, imposing an unexpected burden of care on elderly adults (Knodel and VanLandingham, 2003). According to a study in Zimbabwe by the World Health Organization, the large majority of main caregivers among people 50 years and above were over 60 years old, female and caring for their grandchildren (WHO, 2002), emphasizing the demand on elderly females in particular. As older parents in developing countries commonly expect to rely on adult children for support, the loss of children also affects parents in the long-term. Rather than relying on their children, the elderly are finding themselves caring for children and grandchildren, causing extreme financial strain. The effect in Thailand is somewhat less severe due to basic governmental coverage of treatment costs, excluding expensive antiretroviral treatments (Knodel et al., 2001). However, in Zimbabwe, the elderly who care for orphans are often unable to meet the needs of the household even in the presence of pensions allowance which are too little (WHO, 2002).

Impacts on the health of elderly parents include: physical strain from care-giving and extra work required for needed expenses; potential exposure to opportunistic diseases such as TB (Knodel et al., 2001); and a host of other physical illnesses (WHO, 2002). One study in Tanzania found that the death, primarily from AIDS, of an adult in well-off households lowered the body mass index (a measure of physical well-being) of the elderly in those households prior to the death, partly from a decline and diversion of resources to patients (Ainsworth et al., 2005). Emotional strain also affects parents and families, particularly due to the extended nature of the disease (Knodel et al., 2001). Burnout, stress and worry are common for elderly caregivers as are experiences of abuse, both from outsiders due to AIDS-related stigma and discrimination as well as from their own sick children (WHO, 2002).

Applications of Generalized Linear Mixed Models to Disease Impact Modelling

Generalized Linear Mixed Models (GLMMs), among other modelling strategies, have been applied variedly in different fields of human endeavor in recent years. It all started with Generalized Linear Models (GLMs), introduced in an article by Nelder and Wedderburn (1972), which is also an extension of the traditional Linear Models (LMs), which allows a population mean to be dependent on a linear predictor (through a link function), thereby allowing the response probability to be any member of a family of exponential distributions. Following that, a detailed introduction to GLMs was published by McCullagh and Nelder (1989). After these, many other publications followed with applications of GLMs to different fields of activity Epidemiology (Zuccolo et al., (2005), Kleinman, K., Lazarus, R. and Platt, R., 2014; Duffy, 1989) and Public Health (Das et al., 2004), among others. Aitkin et al., (1989) and Dobson (1990) also published very useful references on applications of GLMs, up to the point where

Haberman and Renshaw (1996) also came up with a publication on the applications of GLMs to actuarial problems.

Generalized Linear Mixed Models, which is an extension of GLMs altered with random effects, in recent times, appear to be taking over from GLMs, in terms of popularity. This follows many other recent publications of GLMMs with applications to Medicine (Burton P.R., 2003; Burton et al., 1999), Public Health (Burton, 2003; Schachterle et al., 2013, etc.) and Epidemiology (Duffy, 1989; Zuccolo et al., 2005; Hunger et al., 2012, etc.), among others. McCulloch and Searle (2001) and Demdenko (2004) are some of the numerous sources of literature on the modelling of binary or count, clustered and longitudinal data. Another very useful literature source on the applications of GLMMs is Antonio and Beirlant (2007) and this happens to be in the field of actuarial statistics.

However, no publication has so far been seen on the applications of GLMMs with respect to the impact of diseases on households. The closest that we have had is in the area of Linear Regression by Pitayanon et al. (1998) in Thailand and Zhang et al. (2012) in China. While the former looked at the economic and social impact of morbidity and mortality of HIV/AIDS on the household and their coping mechanisms, the latter looked at factors associated with per-capita income in AIDS-affected households. They did not look at the relative impact of morbidity and mortality of other diseases on the household using the GLMM. They only limited their studies to Linear Models (LMs).

So far, Eze (2009) is among the few who did some work which is closest to what this study is pursuing. He developed statistical models describing the spatial distribution of the HIV/AIDS epidemic in Nigeria and its associated ecological risk factors, reconstructing the HIV incidence curve and obtained an estimate of the hidden HIV/AIDS population and a short term projection for AIDS incidence and a measure of precision of the estimates. Along the line, he used Binary Logistic Regression to determine the effects of explanatory variables such as sex, age and time period of the test on the test outcome. Mzolo et al. (2009) also did a classical piece titled “Bayesian versus Frequentist Approaches in Risk Determinants of Infectious Diseases”. In their study, they used the GLMM to ascertain the determinants of HIV and TB. Among their fixed effects factors were age, sex, educational qualification, income status, race group, condom use at first sex, and health status, while their random factor was enumeration area. Their findings therefore indicated, among other findings in the significant fixed factor determinants, that given the estimated intra-class correlations for HIV and TB being 0.169 and 0.249 respectively, people in the same enumeration area were more likely to be correlated in their risk of HIV as well as TB, respectively.

In Ghana, several studies have also been conducted to ascertain the socio-economic impacts of HIV/AIDS on households (Bollinger et al., 1999; Kwankye, 2000; Oppong, 2001). However, they fell short of using any Linear Modeling, General Linear Modeling or Generalized Linear Mixed Modeling approach to arrive at their impacts. They only centered on descriptive statistics.

Review of Similar Studies

A review is made of two similarly conducted studies, one in Thailand (Pitayanon et al., 1998) and the other in South Africa (Mzolo et al., 2009).

The Economic Impact of Adult AIDS Deaths on Rural Households in Thailand

"The Impact of AIDS", released in October, 2004 and published by Population Studies Series, (2004), measures and analyses the economic impact of adult AIDS deaths on rural households in Thailand based on a primary survey of rural households in one of the provinces carried out by Pitayanon S., Kongsin S. and Janjareon W. S. (1997). Among other analyses, the survey investigates whether an adult AIDS death differs from a death from other causes in terms of the economic impact on the household.

The main object of this study was to measure and analyse the economic impact of an adult HIV/AIDS-related death on a rural Thai household based on a primary data survey of rural households in Chiangmai province in northern Thailand. The data used in this study were generated from a field-based survey of households with recent experience of HIV-AIDS-related death in five districts of Changmai province. Household selection was based on hospital records of HIV/AIDS-related deaths. A total of 116 households, with recent experience of an HIV/AIDS-related death, were interviewed. The survey also included 100 households where a non-HIV/AIDS-related death had occurred and 108 households where no death had occurred as a control group, making a total of 324 households.

The measurement of the economic impact of HIV/AIDS mortality on households was based on the calculation of direct and indirect costs of death, the investigation of household coping strategies and the determination of the real economic impact of death from HIV/AIDS. In addition to comparative analyses of socioeconomic characteristics of survey households, direct and indirect costs of an HIV/AIDS-related death and Non HIV/AIDS-related death on a household, economic and other socioeconomic impacts of HIV/AIDS-related death and non-HIV/AIDS-related death on households, among others, there was the application of the General Linear Model, in particular linear regression analyses, to investigate whether an HIV/AIDS-related death actually makes a difference to the economic condition of the affected household. In this regard, two key dependent variables explored were household income and household change in consumption.

The socioeconomic factors included in the regression model as the determining factors of household income and household consumption change after death were household size, sex of the deceased, age of the deceased at death, household status of the deceased, cause of death, occupation of the deceased before death, and educational attainment of the deceased, most of which were dummy variables. The simple regression analyses indicated that an adult HIV/AIDS-related death caused a greater negative impact on household income and a larger consumption change to the household than an adult death not HIV/AIDS-related.

The study found that the economic impact of an adult AIDS death is sizeable and significant and the least able to cope with adults AIDS death were the poorest and the least educated households engaged in agricultural work. It was also found that the economic impact of an adult AIDS death was more severe than the impact of death from other causes

Review of Bayesian vrs Frequentist Approaches in Estimating the Risk Determinants of Infectious Diseases

The second study applied a multi-stage disproportionate stratified sampling with sampling frame obtained from the 2005 survey based on a sample of 1000 enumeration areas. They used the GLMM where the linear predictor, $\boldsymbol{\eta}$, contains fixed and random effects such that $\boldsymbol{\eta} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma}$. Here $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are vectors of fixed and random effects respectively while $\boldsymbol{X}$ and $\boldsymbol{Z}$ are design matrices for the fixed and random effects respectively. Here the fixed effects include sex, age, income, race group, education, health status and condom use at first sex. The random effects were the enumeration areas which were assumed to follow a normal distribution centered at zero. The response variable HIV followed a Bernoulli distribution with 1 if infected and 0 otherwise.

The dataset included 16,398 observations. However, after withdrawing records with missing cases, 9,412 observations remained.

In the resulting model, the variance was estimated to be 0.1659 with a standard error of 0.04718. Their result therefore confirmed that males were less likely to be infected with HIV as compared to females and as the age group increased, the chances of being infected with HIV also decreased. The odds of being infected also varied according to racial groupings. The rate of HIV infection also varied according to educational qualification while individuals who were in good health were less likely to be infected with HIV than those in poor health.

Also, those using a condom at first sex were less likely to be infected with HIV than those who did not use a condom at first sex.

With respect to the random effects, there was a positive correlation at the EA level. Their interpretation of the random effect therefore was that any intervention in HIV should consider the EA level effect rather than the individuals.

CHAPTER THREE

REVIEW OF METHODS

Introduction

This chapter presents a review of Generalized Linear Models (GLMs), with a look at the model fit procedures, estimation procedures and inferences in GLMs. Further in this chapter, a review is made of Generalized Linear Mixed Models (GLMMs) with respect to model specification, dwelling on the subject-specific model and the population averaged model, parameter estimation procedures, with a look at the G-side and R-side random effects, as well as model fit statistics, among others. A look is then taken at applications of GLMMs to disease modeling and the PROC GLIMMIX modeling procedure in SAS. At the tail end of the chapter, a presentation is made on how the survey instrument was pre-tested, the sampling strategy employed in the field work, procedures for data collection, as well as how the economic impact measurements were done. Finally, the variables used in the analyses are also presented in this chapter.

Generalized Linear Models

Generalized Linear Models are a broad view of the very well-known General Linear Models that allow the mean of a population response variable to depend on a set of linear predictors through a link function, which could be non-linear in the exponential family (such as binomial, normal or poisson). This permits the probability distribution of the response variable to be any member of the exponential family of distributions.

A Generalized Linear Model consists of three key components described as follows:

For a parameter θ, and a dispersion parameter ϕ, the response Y has a distribution in the exponential family, with density (or probability) function of the form

$$f(y;\theta,\emptyset) = \exp\left\{\int \frac{y-\mu(\theta)}{\phi V(\mu)} d\mu(\theta) + c(y,\phi)\right\}, \qquad (3.1)$$

for a given mean, $\mu(\theta) = \mathbb{E}(Y)$, and variance, $\mathbb{V}(Y) = \phi V(\mu)$ and known bivariate function c. The exponential family is very flexible and can model continuous, binary, or count data.

For some vector of parameters $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$, and covariate $\underline{X_i} = (x_{i_1}, x_{i_2}, \dots, x_{i_p})'$ associated with observation Y_i, and a random sample $Y_{1,\dots,}Yn_{,}$ the linear component, η_i, is defined as

$$\eta_{i} = \boldsymbol{X}_i' \boldsymbol{\beta}, \qquad i = 1, \dots, n, \qquad (3.2)$$

For a linear predictor η_i, a monotonic differentiable link function g describes how the expected response $\mu_i = \mathbb{E}(Y_i)$ is related to the linear predictor η_i by

$$g(\mu_i) = \eta_{i}, \qquad i = 1, \dots, n. \qquad (3.3)$$

The different components of GLMs most commonly used in real life applications can be found in Appendix F.

By way of model interpretation for the logit link, which is defined as the estimated change in the probability of success that is commensurate with one-unit change in the associated predictor, the calculation of odds ratios helps in the interpretation of the predictors.. The associated odds ratio is obtained by

$$\hat{O}_R = \frac{\text{Odds}_{(x_{i+1})}}{\text{Odds}_{(x_i)}} = e^{\hat{\beta}_i},$$

where $\hat{O}_R$ is the odds ratio for the predictor variable in question and $\hat{\beta}_i$ is the coefficient of the corresponding predictor variable in the model.

Maximum Likelihood Estimate of the regression parameters

Consider the parameter $\boldsymbol{\beta}$. The likelihood of $\boldsymbol{\beta}$ is given, in general terms, as

$$L(\boldsymbol{\beta}/\boldsymbol{y}) = f_{\boldsymbol{\beta}}(\boldsymbol{y}/\boldsymbol{\beta}),$$

for a set of n independently and identically distributed (IID) observations, y_1, y_2, ... y_n. For a normal distribution, the likelihood of $\boldsymbol{\beta}$ is therefore given as

$$l(\boldsymbol{\beta}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2$$

The proof is presented in Appendix E.

The log-likelihood function of $\boldsymbol{\beta}$, $\ln L(\boldsymbol{\beta})$, denoted simply by $l(\boldsymbol{\beta})$ is given in general terms (i.e. for any distribution) as

$$l(\boldsymbol{\beta}) = \ln L(\boldsymbol{\beta}) = \sum_{i=1}^{n}\left\{\int \frac{[y_i - \mu_i(\theta)]}{\phi V(\mu_i)} d\mu_i(\theta) + c(y_i, \phi)\right\} \qquad (3.4)$$

The proof is presented in Appendix E. Below are illustrations of log-likelihood functions for certain universally used distributions.

1. Normal:

$$l(\boldsymbol{\beta}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2$$

The proof of the Normal distribution is presented in Appendix E.

2. Poisson:

$$l(\boldsymbol{\mu};\boldsymbol{Y}) = \sum_{i=1}^{n} y_i \ln \mu - n\mu$$

The proof follows from those of the Normal and Poisson, from Appendix E, for the remaining distributions stated below:

3. Gamma:

$$l(\boldsymbol{\beta}) = \sum_{i-1}^{n}\left\{\frac{\omega_i}{\phi}\,\mathrm{In}\left(\frac{\omega_i y_i}{\phi\mu_i}\right) - \frac{\omega_i y_i}{\phi\mu_i} - \ln(y_i) - \ln\left(\Gamma\left(\frac{\omega_i}{\phi}\right)\right)\right\}$$

4. Inverse Gaussian:

$$l(\boldsymbol{\beta}) = \sum_{i=1}^{n} -\frac{1}{2}\left\{\frac{\omega_i (y_i - \mu_i)^2}{y_1 \mu^2 \phi} + In(\frac{\phi y_i^3}{\omega_i}) + \ln(2\pi)\right\}$$

5. Negative Binomial:

$$l(\boldsymbol{\beta}) = \sum_{i=1}^{n}\left\{ y_i \ln(k\mu) - \left(y_i + \frac{1}{k}\right)\ln(1 + k\mu) + \ln\left(\frac{\Gamma\left(y_i + \frac{1}{k}\right)}{\Gamma(y_i + 1)\Gamma\left(\frac{1}{k}\right)}\right)\right\}$$

6. Multinomial:

$$l(\boldsymbol{\beta}) = \sum_{i=1}^{n}\{y_{ij}\ln(\mu_{ij})\}$$

When we maximize the log-likelihood function in equation (3.4) and solve for $\boldsymbol{\beta}$, we obtain the MLE of the regression parameter, $\boldsymbol{\beta}$. Now, taking the derivative of equation (3.4) gives

$$\frac{dl(\boldsymbol{\beta})}{d\boldsymbol{\beta}} = \sum_{i=1}^{n} \frac{dl(\boldsymbol{\beta})}{d\mu_i} \frac{d\mu_i}{d\boldsymbol{\beta}} = \sum_{i=1}^{n} \frac{(y_i - \mu_i)}{\emptyset V(\mu_i)} \frac{d\mu_i}{d\boldsymbol{X'}_i\boldsymbol{\beta}} \frac{d\boldsymbol{X'}_i\boldsymbol{\beta}}{d\boldsymbol{\beta}}$$

and

$$\frac{d\mu_i}{d\boldsymbol{X'}_i\boldsymbol{\beta}} = \frac{dg^{-1}(\boldsymbol{X'}_i\boldsymbol{\beta})}{d\boldsymbol{X'}_i\boldsymbol{\beta}} = \frac{1}{g(\mu_i)}$$

Therefore,

$$\frac{dl(\boldsymbol{\beta})}{d\boldsymbol{\beta}} = \sum_{i=1}^{n} \frac{(y_i - \mu_i)}{\emptyset V(\mu_i)} \frac{1}{g'(\mu_i)} \boldsymbol{X'}_i \tag{3.5}$$

For a normal distribution $g'(\mu_i) = 1$, and V $(\mu_i) = 1$ for all i.

To maximize, we put

$$\frac{dl(\boldsymbol{\beta})}{d\boldsymbol{\beta}} = 0$$

giving

$$\sum_{i=1}^{n} X_i(y_i - \boldsymbol{X'}_i\boldsymbol{\beta}) = 0$$

This closed form of solution does not exist in other forms of exponential family cases for this system of p equations. For such instances, to obtain the maximum likelihood estimator (MLE) of model parameters numerically, we employ iterative algorithms such as the Newton-Raphson or Fisher scoring methods.

The Newton-Raphson Method of Parameter Estimation

With the Newton-Raphson method, sequential approximations are provided to the root $\boldsymbol{\beta}$ of Equation (3.5). On the *r*th iteration, the parameter estimate $\widehat{\boldsymbol{\beta}}_r$ is updated by the algorithm

$$\widehat{\boldsymbol{\beta}}_{r+1} = \widehat{\boldsymbol{\beta}}_r - \boldsymbol{H}^{-\mathbf{1}}{}_{\boldsymbol{s}} \qquad\qquad r = 1, 2, \ldots$$

where **H** is the Hessian matrix, and $\boldsymbol{s}$ is the gradient vector of the log-likelihood function, which are both evaluated at the current value of the parameter estimate and are given by

$$\boldsymbol{s} = \sum_i \frac{\omega_i (y_i - \mu)\boldsymbol{x_i}}{V(\mu_i) g'(\mu) \emptyset}$$

Where V is the variance function and $x_{\boldsymbol{i}}$ is the transpose of the *i*th row of the design matrix **X**. The matrix **H** is given as

$$\boldsymbol{H} = -\boldsymbol{X}'\boldsymbol{W}_{\mathbf{0}}\boldsymbol{X}$$

where **X** is the design matrix, $\boldsymbol{x}_i$ is the transpose of the *i*th row of **X,** and *V* is the variance function. The matrix $\mathbf{W}_o$ is a diagonal one. $\mathbf{W}_o$ has its *i*th diagonal element equal to

$$\omega_{oi} = \omega_{ei} + \omega_i (y_i - \mu_i) \frac{V(\mu) g''(\mu) + V'(\mu_i) g'(\mu_i)}{[V(\mu_i)]^2 [g'(\mu_i)]^3 \phi},$$

where

$$\omega_{ei} = \frac{\omega_i}{\phi V(\mu_i)[g'(\mu_i)]^2},$$

where, ω_i is a known weight for each observation. When the weight is not known, we put ω_i = 1 for each observation.

The prime components denote derivatives of g and V with respect to μ. Here, the negative of **H** is called the observed information matrix. The expected value of $\mathbf{W}_o$ is a diagonal matrix $\mathbf{W}_e$ with diagonal values ω_{ei}. Now, when $\mathbf{W}_o$ is replaced with $\mathbf{W}_e$, the negative of **H** is called the expected information matrix. $\mathbf{W}_e$ is the weight matrix for Fisher's scoring method.

The GLM theory was developed for dependent observations in the exponential family of distributions. However, the theory together with its numerical algorithm extends to other distributions outside the exponential family.

Asymptotic properties of the General Linear Model MLE

When the number of observations n approaches infinity, the MLE $\widehat{\boldsymbol{\beta}}$ of the GLM parameters exhibit some asymptotic properties. Hence, $\boldsymbol{\beta}$ becomes an asymptotically unbiased and consistent estimator of $\boldsymbol{\beta}$. Thus

$$V\left(\underline{\boldsymbol{\beta}}\right) \rightarrow \sum = \boldsymbol{H}^{-1} \text{ as } n \rightarrow \infty$$

and

$$\boldsymbol{H} = -\boldsymbol{X}'\boldsymbol{W}_{\mathbf{0}}\boldsymbol{X}$$

is the Hessian matrix, while

$$W_o = diag\ (w_{o1}, \ldots, w_{on})$$

is a diagonal weight matrix with *i-th* element

$$\omega_{oi} = \frac{\omega_i}{\phi V(\mu_i)(g'(\mu_i))^2} + \omega_i (y_i - \mu_i) \frac{V(\mu_i) g''(\mu_i) + V'(\mu_i) g'(\mu_i)}{(V(\mu_i))^2 (g'(\mu))^3 \phi},$$

for known weights ω_i and covariate matrix $\boldsymbol{X} = (\widehat{\boldsymbol{X}}_1, \ldots, \widehat{\boldsymbol{X}}_n)$

The distribution of $\boldsymbol{\beta}$ is given as

$$\widehat{\boldsymbol{\beta}} \overset{d}{\rightarrow} N(\beta, (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\emptyset)$$

thus it converges in distribution [Fahrmeir and Kaufman, 1985].

For a finite sample, the MLE, $\widehat{\boldsymbol{\beta}}$, of $\boldsymbol{\beta}$ is most commonly biased. Thus its mean square error (MSE),

$$\text{MSE}\ (\widehat{\boldsymbol{\beta}}) = V(\widehat{\boldsymbol{\beta}}) + \text{bias}(\widehat{\boldsymbol{\beta}})$$

plays an important role, where $\text{bias}(\widehat{\boldsymbol{\beta}}) = E(\widehat{\boldsymbol{\beta}}) - \boldsymbol{\beta}$

Wald inference is employed in testing for the significance of each of the parameters in the hypotheses $H_0: \beta_i = 0$ against $H_0: \beta_i \neq 0$ producing z-statistics and accompanying p-values, comparable to the situation of the linear regression where t-test is applied, for each of the link functions.

According to McCullagh and Nelder (1989), Myers et al., (2010) and Hosmer and Lemeshow (2000), there are three statistics that are used in determining the adequacy of the ensuing model by goodness-of-fit tests. These are the Pearson's χ^2, the Deviance and Hosmer-Lemeshow values.

Generalized Linear Mixed Models (GLMMs)

Emanating from the Linear Model (LM) is the General Linear Model out of which is also obtained the Generalized Linear Model (GLM), where the non-linear link function is employed as the response variable, possessing the fundamental discrete and continuous distributions. The GLM is fitted using either the Ordinary Least Squares (OLS) or the Maximum Likelihood Estimation (MLE) methods.

The classical GLM is generally written as the sum of two parts (a fixed component $\boldsymbol{X\beta}$, which is a linear function of the independent coefficients, and a random noise, $\boldsymbol{\varepsilon}$, also called the random error component of $\boldsymbol{Y}$, and is written as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Where $\boldsymbol{\varepsilon}$**,** also called the random or stochastic component of $\boldsymbol{Y}$ is $NID\ (\mu_i, \sigma^2)$. This is sometimes called the "Error Structure" or the "Response Distribution" (Gill, 2001). Thus, each component of $\boldsymbol{Y}$ is assumed to differ in mean μ_i from each other but all have common variance σ^2. The k covariates are assumed to combine to yield the "linear predictor" $\boldsymbol{\eta} = \boldsymbol{X}.\boldsymbol{\beta}$ (or the systematic component). The random and systematic components are assumed to be related through a link function $E(\underline{Y}) = \boldsymbol{\mu} = \boldsymbol{\eta}$ (for the linear model, the link function is the identity function).

Hence,

$$g(\mu) = \boldsymbol{\eta} = \boldsymbol{X\beta}$$

The generalized linear mixed model (GLMM) is an extension of the GLM. The GLMM is a generalized linear model made up of fixed and random components. Thus, the GLMM is an extension of the GLM complicated by random effects. It is used for modeling binary or count, clustered and longitudinal data.

Among the assumptions of the GLMs is the independence of data. This assumption therefore suggests that the underlying study design is completely randomized. However, according to Robinson et al. (2004), in practical applications, this assumption is very commonly violated. A typical example is in the collection of longitudinal data where a subject is studied over time. The resulting data is correlated, thereby violating the assumption of independence.

Model Specification for GLMMs

According to Robinson et al. (2004), GLMMs stretch the Generalized Linear Model (GLM) such that it accounts for correlations that are present in the random effects. Agresti (2002) also states that the random effects model can also take care of methods of addressing missing data

A GLMM therefore may consist of a number of components. For cluster data Y_{ij}, $i = 1, \ldots, n$ and $j = 1, \ldots, n_i$, assumed conditionally independent given the random effects $\mathbf{U}_1, \ldots, \mathbf{U}_n$, consider the following distribution:

$$f(y_{ij}|\boldsymbol{u_i}, \theta, \emptyset) = \exp\left\{\frac{[y_{ij}\theta_{ij} - b(\theta_{ij})]}{\phi} + c(y_{ij}, \phi)\right\} \quad (3.6)$$

where

$$\boldsymbol{u_i} = (\boldsymbol{u}_{i1}, \dots, \boldsymbol{u}_{ik})$$

are variates from normally distributed *k*-dimensional random vectors,

$$\boldsymbol{U_i} \sim \boldsymbol{N}\ (\boldsymbol{0}, \boldsymbol{D})$$

where **D** is the variance-covariance matrix of $\hat{\theta}$ and

$$\mu_{ij} = E[Y_{ij}|U_i - u_i] = b'(\theta_{ij})$$

The variance of the observations, conditional on the random effects, is given by

$$\mathrm{var}[Y_{ij}|\boldsymbol{U}_i] = \boldsymbol{A}_i^{1/2}\boldsymbol{R}_i\boldsymbol{A}_i^{1/2}$$

which is the variance functions of the model, which express the variance of a response Y_{ij} as a function of its means $u_{ij.}$ The diagonal matrix $\mathbf{A}_i$ then conditions the variance function of the model. The random effects of the model have a variance-covariance matrix represented by $\mathbf{R}_i$ which is the variance-covariance matrix of the random effects.

The linear mixed effects model is defined as

$$\eta_{ij} = \boldsymbol{X'}_{ij}\boldsymbol{\beta} + \boldsymbol{Z'}_{ij}\boldsymbol{\gamma} \qquad i = 1, \dots, n \quad j = 1, \dots, n \quad (3.7)$$

For the fixed effects parameter vector $\boldsymbol{\beta} = (\beta_1,..., \beta_p)'$ and random effects vector $\boldsymbol{\gamma} = (\gamma_{i1}, \dots, \gamma_{ik})'$

It should be noted that $\boldsymbol{X}_{ij} = (x_{ij1}, \dots, x_{ijp})$ and $\boldsymbol{Z}_{ij} = (z_{ij1}, \dots, z_{ijp})$ are covariates, and the link function g, is given by

$$g(u_{ij}) = \eta_{ij} \qquad i = 1, \dots, n \quad j = 1, \dots, n \tag{3.8}$$

The principle of likelihood is mostly applied to estimation methods for parameters $\underline{\beta}$ and $\underline{u}_i$ of GLMMs and for such, the estimates of the parameters are obtained mostly by quantitative methods. A momentary evaluation of such quantitative methods such as pseudo-likelihood is provided by Antonio and Beirlant (2007). They further provide the Gauss-Hermite quadrature and Bayesian methods.

Four types of procedure are outlined by Demidenko (2004) for obtaining the GLMM and these include: (a) maximum likelihood with numerical quadrature, (b) penalized quasi-likelihood (PQL), (c) specific methods in conjunction with a Laplace approximation or a generalized estimating equation (GEE) approach, and (d) Monte Carlo methods for integral of likelihood approximations.

The log-likelihood for the GLMM defined in equations (3.6) through (3.8) takes the following form:

$$l(\underline{\beta}, D) = -\frac{nk}{2}\ln(2\pi) - \frac{n}{2}|D| + \sum_{i=1}^{n} \ln \int_{R^k} \exp\left[l_i(\underline{\beta}, v) - \frac{1}{2} v' D^{-1} v \right] dv, \tag{3.9}$$

Where

$$l_i(\boldsymbol{\beta}, \boldsymbol{v}_i) = \sum_{j=1}^{n_i} \left[\left(\boldsymbol{X'}_{ij}\boldsymbol{\beta} + \boldsymbol{T'}_{\boldsymbol{ij}}\boldsymbol{v}_i \right) y_{ij} - b\left(\boldsymbol{X'}_{ij}\boldsymbol{\beta} + \boldsymbol{T'}_{\boldsymbol{ij}}\boldsymbol{v}_i \right) \right] \qquad (3.10)$$

is the ***i***-th conditional log-likelihood.

Two types of numerical algorithms exist to solve for (3.9), the first of which is based on the Taylor series, and are therefore referred to as linearization methods. The expansion of these Taylor series yields an approximate model pseudo-data where there are fewer non-linear components.

According to some criteria, this linear calculation can be done repetitively until convergence is achieved. By way of examples, some procedures are provided based on Taylor series for clustered data by Schaben-berger and Gregoire (1996).

The linearization fitting techniques are a result of two iterations. The first is the GLMM which is approximated by a linear model based on current values of the covariance parameter estimates. The end result which is a linear mixed model is then fitted, forming an iterative process.

During convergence, the newly estimated parameters are then used in updating the linearization, resulting in a new linear mixed model. The iteration process ends when parameter estimates, for successive fits of the linear mixed model, change only within a specified margin.

The second type of algorithm is based on integral approximations, where the numeral optimization follows the approximation of the log-likelihood of the GLMM. Approximation techniques such as Markov Chain Monte Carlo, Monte

Carlo integration and Laplace and quadrature methods are among the various approximation techniques that exist.

The Subject-Specific Model

Myers et al. (2010) indicate that the subject-specific model become more relevant in repeated measures studies where measurements are taken on subjects longitudinally (or over a length of time). The resulting models therefore yield estimates of the mean based on the levels of the random effects or components of the model.

According to Myers et al. (2010), Random effects GLMs are given by

$$y = \mu + \varepsilon$$

given that

$$g(\mu) = \boldsymbol{X}\beta + \boldsymbol{Z}\gamma$$

Assuming γ and ε are independent and g is the appropriate link function, $y|\gamma$ has a distribution of the exponential family and each of the random effects ~ $N(0, G_i)$ and the G_i are the same for each cluster, then the conditional mean of the i^{th} cluster is given by

$$E\left(y_{n_i}|\gamma_i\right) = g^{-1}(\eta_i) = g^{-1}(\boldsymbol{X_i\beta} + \boldsymbol{Z_i}\gamma_i)$$

where y_{n_i} is the i^{th} cluster's response vector

η_i, the linear predictor,

$\boldsymbol{X_i}$, the $(n_{\boldsymbol{i}} \times p)$ matrix of fixed effect models terms associated with the i^{th} cluster, and

$\boldsymbol{\beta}$ is the corresponding (p $\times$ 1) vector of fixed effect regression coefficients.

and the i^{th} cluster has $n_{\boldsymbol{i}}$ observations.

The random effect component of the model, γ_i, becomes the (q $\times$ 1) vector of random factor levels associated with the i^{th} cluster and $\boldsymbol{Z_i}$ is the corresponding matrix of predictors of the i^{th} cluster

The Population Averaged Model

According to Myers et al. (2010), on occasions when the interest of the modeling is not in more specific levels but rather in the estimation of more general trends across the entire population of random effects, then a population averaged model is more applicable. Meanwhile, a more common way of estimating the marginal mean using the batch-specific models is to make $\hat{\gamma} = 0$ given that $E(\gamma) = \mathbf{0}$. This way, the estimate of the marginal mean will certainly be different from what pertains when the population-averaged approach is used where the conditional effects are generally larger than the marginal effects. According to Agresti (2002), this also most commonly results in similar significance of the effects.

More difficult to obtain in practice is the marginal mean. This could be attributed to non-linearity in the GLMMs. Frequently, a way around this is by approximations through linearizing of the conditional mean. Performing a first-order Taylor series expansion on $E(\boldsymbol{\eta}) = \boldsymbol{X\beta}$ gives the linear form of the unconditional mean as $E(y) = E[E(y|\gamma)] \approx g^{-1}(x\beta)$.

Myers et al. (2010) however indicate that when the variance components associated with δ approach 0, the approximation of the linear link function is more exact.

One significant difference between the population averaged model and the batch-specific one is the fact that whereas the former requires the definition of a covariance structure for the error term, the latter does not. According to Robinson et al. (2004), the correlation matrix for a split-plot design has a compound symmetric structure. In a time series study (i.e. following subjects longitudinally and collecting repeated data on them), the random effect $\boldsymbol{R}$ assumes a first order auto-regressive structure.

In modeling, when the prediction of an average across all subjects, such as in batches (random effects) is of interest, it is better to model the unconditional expectation of the response than the conditional expectation when the focus is on examining the application of both the population-average and batch-specific models to a split plot study design (Robinson et al., 2004). It is also worth noting, from Robinson et al. (2004), that though the population-averaged model appears more likeable than the batch-specific one, the former is more heavily dependent on the assumption that the group of random subjects (or clusters in this case) is a true representation of the whole.

Linearization-Based Pseudo-Likelihood Estimation

Unlike in the case of GLMs, where the independence of the underlying data makes the log-likelihood well-defined and the objective function for estimating the parameters is simple to construct, it may not be possible to compute the objective function in the case of the GLMMs (SAS, 2010). According to SAS (2010), several problems in GLMMs must be solved in order to compute an objective function. These problems have been listed as the situation where

- no valid joint distribution can be constructed either in general or for a particular set of parameter values;
- the dependency between mean and variance for non-normal data places constraints on the possible correlation models that simultaneously yield valid joint distributions and desired conditional distributions, and
- even if the joint distribution is feasible mathematically, it still can be out of reach computationally when data are independent or conditional on the random effects from the joint distribution. However, numerical integration is practical only when the number of random effects is small and when the data has a clustered (subject) structure.

To get around this problem, SAS (2010) has suggested two alternative approaches;

- Approximation of the objective function, and
- Approximation of the model.

Furthermore, according to SAS (2010), techniques such as Laplace methods, quadrature methods, Monte Carlo integration, and Marcov Chain Monte Carlo methods are used by integral approximation methods in approximating the log likelihood of the GLMM by using the approximated function in numerical optimization. This approach has an advantage of providing an actual objective function for optimization.

With the use of expansions to approximate the model based on pseudo data with fewer non-linear components, linearization methods are used to approximate the model. The iteration process is such that the GLMM is approximated by a linear mixed model based on current values of the covariance parameter estimates, producing a linear mixed model which in turn is fit using an iterative process. The new parameter estimates are then used to update the linearization during convergence. This cycle continues until the parameter estimates between consecutive linear mixed model fits change within a specified tolerance or converge.

The linearization based method of GLMM modeling does not use a true objective function for the overall optimization process, though it handles models with correlated error. However, while that is the case, where there is a large number of random effects, crossed random effects and multiple types of subjects, it does not use the true objective function for the overall optimization process. This could therefore make the resulting estimates of the covariance parameters potentially biased, especially for binary data.

The linearization based method has a disadvantage because of the absence of a true objective function for the overall optimization process, and this could potentially lead to biased estimates, especially when the number of observations is small for binary data (SAS, 2010).

Now, for the pseudo model, from Equations (3.7) and (3.8) and the SAS manual (SAS, 2010), we have

$$E[\boldsymbol{Y} \mid \gamma] = g^{-1}(\boldsymbol{X\beta} + \boldsymbol{Z}\gamma) = g^{-1}(\boldsymbol{\eta}) = \boldsymbol{\mu},$$

For $\gamma \sim N(0, \boldsymbol{G})$ and $\text{var}[\boldsymbol{Y}|\gamma] = \boldsymbol{A}^{1/2}\boldsymbol{R}\boldsymbol{A}^{1/2}$, given that the vector of responses $\boldsymbol{Y} = (Y_1 \boldsymbol{.} Y_2, \dots, Y_n)'$ are independent,

$\boldsymbol{X} = (\boldsymbol{X'}_1,\ \boldsymbol{X'}_{\boldsymbol{2}}, \dots, \boldsymbol{X'}_n)$ is a matrix of covariates,

$\boldsymbol{Z} = (\boldsymbol{Z'}_1,\ \boldsymbol{Z'}_{\boldsymbol{2}}, \dots, \boldsymbol{Z'}_n)$ is a covariate matrix of random effects,

$\boldsymbol{\beta} = (\beta_1,\ \beta_2, \dots, \beta_n)'$ is a vector of fixed effects and

$\boldsymbol{\gamma} = (\gamma_1,\ \gamma_2, \dots, \gamma_n)'$ is a vector of random effects.

The first-order Taylor series of $\boldsymbol{\mu}$ about $\widetilde{\boldsymbol{\beta}}$ and $\widetilde{\boldsymbol{\gamma}}$ yields

$$g^{-1}(\boldsymbol{\eta}) = g^{-1}(\widetilde{\boldsymbol{\eta}}) + \widetilde{\boldsymbol{\Delta}}\widetilde{\mathbf{X}}\left(\boldsymbol{\beta} - \widetilde{\boldsymbol{\beta}}\right) + \widetilde{\boldsymbol{\Delta}}\widetilde{\mathbf{Z}}(\boldsymbol{\gamma} - \widetilde{\boldsymbol{\gamma}}) \tag{3.11}$$

where $\widetilde{\Delta} = \left(\frac{\partial g^{-1}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}\right)_{\widetilde{\boldsymbol{\beta}},\widetilde{\boldsymbol{\gamma}}}$ is a diagonal matrix of derivatives of the conditional mean when it is evaluated at the expansion locus (Wolfinger and O'Connell, 1993). This can further be presented as

$$\widetilde{\boldsymbol{\Delta}}^{\boldsymbol{-1}}\left(\boldsymbol{\mu} - g^{-1}(\widetilde{\boldsymbol{\eta}})\right) + \boldsymbol{X}\widetilde{\boldsymbol{\beta}} + \boldsymbol{Z}\tilde{\gamma} = \boldsymbol{X\beta} + \boldsymbol{Z}\gamma \tag{3.12}$$

The left-hand side of Equation (3.12) represents the expected value of

$$\widetilde{\boldsymbol{\Delta}}^{-\mathbf{1}}\left(\boldsymbol{Y}-g^{-1}(\widetilde{\boldsymbol{\eta}})\right)+\boldsymbol{X}\widetilde{\boldsymbol{\beta}}+\boldsymbol{Z}\tilde{\gamma} \equiv \boldsymbol{P} \tag{3.13}$$

conditional on γ, and the variance-covariance matrix

$$\operatorname{var}[P|\gamma]=\widetilde{\boldsymbol{\Delta}}^{-\mathbf{1}}\boldsymbol{A}^{1/2}R\boldsymbol{A}^{\mathbf{1/2}}\widetilde{\boldsymbol{\Delta}}^{-\mathbf{1}}\,. \tag{3.14}$$

Hence,

$$\boldsymbol{P}=\boldsymbol{X\beta}+\boldsymbol{Z\gamma}+\ \boldsymbol{\varepsilon} \tag{3.15}$$

can be considered as a linear mixed effects model with a pseudo-response $\boldsymbol{P}$, fixed effects $\boldsymbol{\beta}$, random effects $\boldsymbol{\gamma}$, and $\operatorname{var}[\boldsymbol{\varepsilon}]=\operatorname{var}[\boldsymbol{P}|\gamma]$.

E equation (3.15) is also the linear mixed pseudo model.

Let

$$\boldsymbol{V}(\boldsymbol{\theta})=\boldsymbol{ZGZ}'+\ \widetilde{\boldsymbol{\Delta}}\boldsymbol{A}^{1/2}\boldsymbol{R}\boldsymbol{A}^{1/2}\widetilde{\boldsymbol{\Delta}}^{-1} \tag{3.16}$$

be the marginal variance function in the linear mixed pseudo-model. Then $\boldsymbol{\theta}$ is the $(q\ \times\ 1)$ parameter vector containing all unknowns in $\boldsymbol{G}$ and **R.** Assume further that the distribution of $\boldsymbol{P}$ is known. Then an objective function can be defined based on this linearized model. For a maximum pseudo log-likelihood, $l(\boldsymbol{\theta},\boldsymbol{p})$ for all $\boldsymbol{\theta}$ and $\boldsymbol{p}$,

$$l(\boldsymbol{\theta},\boldsymbol{P})=\ -\frac{1}{2}\left[\sum_{i=1}^{n}\ln|V(\boldsymbol{\theta}_i)|-\ \sum_{i=1}^{n}\boldsymbol{r}'_iV(\boldsymbol{\theta}_i)^{-1}\boldsymbol{r}_i-f\,\ln(2\pi)\right] \tag{3.17}$$

Where

$$\boldsymbol{r_i} = \boldsymbol{P}_i - \boldsymbol{X}_i\left(\sum_{\boldsymbol{j}=1}^{\boldsymbol{n}} \boldsymbol{X}'_j V(\boldsymbol{\theta}_j)\boldsymbol{X}_j\right)^{\boldsymbol{-1}}\left(\sum_{\boldsymbol{j}=1}^{\boldsymbol{n}} \boldsymbol{X}'_j V(\boldsymbol{\theta}_j)\boldsymbol{P}_j\right)$$

and f denotes the sum of the frequencies used. During convergence, the estimates

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{X}'_i V(\boldsymbol{\theta}_i)\boldsymbol{X}_i\right)^{-1}\left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{X}'_i V(\boldsymbol{\theta}_i)^{-1}\boldsymbol{P}_i\right) \tag{3.18}$$

and

$$\widehat{\boldsymbol{\gamma}}_i = \widehat{D}T'_i V(\hat{\theta}_i)^{-1}(\boldsymbol{P}_i - \boldsymbol{X}_i\widehat{\boldsymbol{\beta}}) \tag{3.19}$$

From the above, the pseudo response and error weights of the linearized model are computed and the objective function is minimized again until the relative change between parameter estimates at two successive iterations is sufficiently small.

The G- and R-side Random Effects

From the standard generalized linear mixed model,

$$\boldsymbol{\eta} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{\gamma} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{G})$, $\boldsymbol{\varepsilon} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{R})$ and Cov $[\boldsymbol{\gamma}, \boldsymbol{\varepsilon}] = 0$

Here the matrices $\boldsymbol{G}$ and $\boldsymbol{R}$ are covariance matrices for the random effects and the random errors respectively. A *G-side* random effect in a mixed model is an element of $\boldsymbol{\gamma}$, and its variance is expressed through an element in $\boldsymbol{G}$, whereas an *R-side* random variable is an element of $\boldsymbol{\varepsilon}$ and its variance is an element of $\boldsymbol{R}$ (SAS, 2010). Furthermore, the *G-side* random effect is inside the link function (and hence the linear predictor) making it easier to interpret. The *R-side* effect on the other hand applies to the covariance matrix on $\boldsymbol{\varepsilon}$ and is outside the link function, for which reason, it is difficult to interpret. Various literature caution that interpretation must be done with extra care (Toy et al., 2011; Verbeke, 2006; SAS, 2010).

Model Fit Tests for GLMMs

According to SAS, the GLIMMIX procedure estimates the parameters for a model containing random effects by default. This is done by applying pseudo-likelihood techniques as in Wolfinger and O'Connell (1993), and Breslow and Clayton (1993). SAS (2010) therefore uses the penalized quasi-likelihood (PQL) method, which is just an approximation, by default. The procedure involves a series of optimizations obtained through iterative estimation methods based on linearizations (using Taylor series expansions). The procedure is such that after each optimization, a new pseudo-model is constructed for the mean response. All the fit statistics (AIC, BIC, etc.) that SAS reports are calculated from the likelihood of the final "pseudo" model, thus the term "pseudo-likelihood" (as will be seen in Chapter Five of this thesis, SAS posts by default Pseudo-AIC, Pseudo-BIC, among others).

The Pseudo-Likelihood concept is therefore applied when the likelihood function is inflexible, but the likelihood of a related but simpler model is available. Hence pseudo likelihood techniques make distributional assumptions to obtain a pseudo-model.

PROC GLIMMIX Procedure in SAS

PROC GLIMMIX is a relatively new SAS procedure for fitting the Generalized Linear Mixed Model (GLMM) in SAS®, although it has been available as a macro for some time (Arrandale, V., 2006). Proc Glimmix is a fast, flexible procedure capable of running linear models (fixed effects), generalized linear models (fixed effects), linear mixed models (fixed and random effects) as well as generalized linear mixed models (fixed and random effects) (Arrandale, V., 2006). It fits statistical models to data with correlations or non-constant variability in situations where the response is not necessarily normally distributed. Generally, it is presented as in Appendix C.

The PROC GLIMMIX and MODEL statements are required, and the MODEL statement must appear after the CLASS statement, if a CLASS statement is included in the syntax. The MODEL statement specifies the fixed effects (the **X** matrix) whereas the first RANDOM statement (the **Z**-matrix), which can be presented as many times as needed in the same model syntax, is used to specify the G-side random effects while the second specifies the R-side random effect. The METHOD option specifies the estimation technique.

There are numerous estimation techniques in SAS. However, only four (MMPL, MSPL, RMPL and RSPL) will be considered in this part of the thesis and two (MMPL and RMPL) will be applied to real life data in latter chapters. The first letter (either “M” or “R”) indicates whether estimation is based on Maximum (M) Likelihood or Residual (R) Likelihood. The second letter (either “M” or “S”) identifies the expansion locus for the underlying approximation, which is either the vector of random effects solution (S) or the mean (M) of the random effects. The expansions are also referred to as Subject-specific (S) or Marginal (M). The last two abbreviations “PL” identifies the method as a Pseudo-Likelihood technique.

Presented in Appendix C (SAS Code B) is a sample of the mixed model, in SAS® procedure, used in this thesis to estimate the determining factors of the likelihood of reallocation of household members’ time as a result of illness and/or death.

When the command “IC=PQ” is included in the PROC GLIMMIX statement, pseudo-AIC and pseudo-BIC values, among others, are included in the output fit statistic table. To determine the better model in terms of fit, both pseudo-AIC and pseudo-BIC must simultaneously post a smaller value for the model with the better fit (Arrandale, 2006).

When ‘Solution’ or ‘s’ is included in the ‘Model’ statement, the fixed effects parameter estimates are produced in the output. However, when it is included in the ‘Residual’ statement, the solution to the random effects (i.e. the estimates of the coefficients, their standard errors, their degrees of freedom, their t-statistic values as well as their p-values) is produced. Also, ‘dist=’ is included in the model statement to specify the distribution of the outcome. The command ‘link’ is specified in the same models statement to specify the link function and ‘or’ is included in the model statement to provide the odds ratios for the fixed effects.

There are families of functions from which outcome variables, distributions and link functions that are employed in the construction of GLMMs using POC GLIMMIX (Table 2). The PROC GLIMMIX statement used in this study, a sample of which is presented in this section, used a binary outcome, a binary distribution and a logit link function. Table 2 provides some functions used in PROC GLIMMIX

Table 2: Functions Used in PROC GLIMMIX

Outcome	**Distribution**	**Link Function**
Beta	Beta	Logit
Binary	Binary	Logit
Binomial	Binomial	Logit
Exponential	Exponential	Log
Gamma	Gamma	Log
Gausssian	Normal	Identity
Geometric	Inverse Gaussian	Inverse Squared
Lognormal	Lognormal	Identity
Multinomial	Multinomial	Cumulative Logit
Negative Binomial	Negative Binomial	Log
Poisson	Poisson	Log
Tcentral	T	Identity

Pre-Testing of the Instrument

The research instrument was pre-tested by administering it to a few individuals. This was done to ascertain its completeness, to establish its reliability and validity. Its content was then validated to remove unclear and ambiguous items and others reformulated. After this, the instrument was pre-tested in the field. The result from the pre-testing was used to improve upon the instrument.

For instance ambiguous questions which were not answered were reframed. Those which did not elicit the required response were also reformulated.

Sampling Strategy

The sampling strategy covers the determination of sample size for the study, the selection of participants for the study and the data collection procedures.

Determination of Sample Size

Given that the study aimed at covering three different populations, three different sample sizes were calculated using the prevalence (proportion) method of sample size calculation. This is because the prevalence of HIV/AIDS is assumed not to be the same as that of non-HIV/AIDS diseases.

Given the Type I and Type II errors as α and β respectively, the sample size n is given as

$$n = \frac{\left(Z_{\alpha/2} + Z_{1-\beta}\right)^2 p(1-p)}{d^2} \tag{3.20}$$

In this study, for No illness/deaths households, we take $\alpha = 5\%$, $(1-\beta) = 80\%$

$p = 50\%$ is the prevalence of ‘No illnesses/deaths’ in Ghana

$d = 10\%$ is the desired level of precision

$Z_{\alpha/2}$ is the critical value for the standard Normal Distribution at α =5%

Hence $n = 198$. Adjusting for a 96% response rate (DHS, 2008), the adjusted sample size n_{adj} is given by $n_{adj} = \frac{n}{\text{resp rate}}$. Hence the adjusted sample size for No illness/deaths households is $n_{adj} = 207$.

The same sample size was used for 'Other Illnesses/Deaths' Households.

For 'HIV/AIDS' Households, we have in Equation (3.20), $\alpha = 5\%$, $1 - \beta = 0.85$, $p = 2.3\%$, which is the 2005 prevalence of HIV in Ghana (WHO, 2007), $d = 10\%$ is the desired level of precision and $Z_{\alpha/2}$is the critical value of the standard normal distribution at $\alpha = 5\%$.

Hence $n = 160$. Adjusting for a 96% response rate, we have $n = 167$. Therefore, a total sample size of 167+207+207=581 households was sampled in all.

It must be noted here that a single sample size could have been computed before splitting it into case-control (using the proportion relationship). However, because HIV/AIDS has its own prevalence which is different from non-HIV/AIDS morbidity, the various case-control sample sizes were computed separately and then merged into one single sample.

Selection of Samples

The HIV/AIDS households were sampled from the National Association of Persons Living with AIDS (NAP+) membership. The NAP+ has smaller units within almost all communities, referred to as 'cells'. The list of all cells in the Greater Accra Region of Ghana was obtained and a simple random sample of six members per cell, including the cell leaders, was drawn. They answered questions on behalf of their respective households. They were interviewed by trained interviewers recruited and trained for the purpose. The remaining two categories of households were interviewed two weeks earlier by the same trained interviewers. With respect to those households, interviewers used simple random sampling within the respective communities to locate a house. Once the house was located, if there was more than one household in that house, interviewers with the help of the researcher, sampled from the number of households available in that household which were eligible for interviewing.

The study targeted a total of 581 respondents (a respondent per household) from the three target populations (167 from 'HIV/AIDS' households, 207 from 'Other Illnesses/Deaths' households and another 207 from 'No Illnesses/Deaths' households). However, at the end of the data collection exercise, 181 HIV/AIDS households were interviewed through NAP+ in the Greater Accra Region (each respondent representing a household), 207 other households in the Greater Accra Region that had experienced a recent (the past one year) non-HIV/AIDS morbidity and/or mortality (Other illnesses/deaths), and 213 different other households which had not experienced any recent (in the past three months) morbidity and/or mortality, respectively, all giving a total sample size of 601. Given that the computed sample size was the minimum required, the final sample of 601 realized from the field was acceptable.

Data Collection Procedures

Data collection took place from December 2008 to January 2009. The respondent was the head of household who responded on behalf of the household. With respect to the 'Other illnesses/death' and 'No illness/death' target groups, interviewers were recruited, trained and despatched to selected communities. Where the selected household was discovered not to have experienced any illness or death in the three months preceding the interview, it was interviewed as a ‘No illness/No death’ household. However, where the selected household was discovered to have had some illness(es) and/or deaths other than that due to HIV/AIDS in the year preceding the interview, it was interviewed as an ‘Other illnesses/deaths’ household. The HIV/AIDS aspect of the study was based on respondents who are members of a PLWHA group in the country called NAP+.

With the assistance of the Ghana AIDS Commission, members of the association were located through an arrangement with their national executives, led by their national President. In this arrangement cell leaders were asked to present a maximum of six members, who had consented to be part of the study from their respective cells, from the very communities where the ‘Other illnesses/deaths’ as well as ‘No illness/death’ households had been interviewed, to their national headquarters. Where a member of the NAP+ was sampled, information on the entire household of that member was collected and where two or more members were sampled and found to belong to the same household, other members belonging to different households were sampled in their place.

The consent of respondents was duly sought before being included in the study and their welfare considered in various stages of the data collection process. The interviews were conducted at a neutral location away from respondents’ homes, since most members of their household did not know about their HIV status. This procedure was adopted upon the request of the respondents.

Accordingly, all the sampled members of NAP+ were invited to their national headquarters where they were interviewed after which they were provided with some incentives.

Economic Impact Measurements

The economic impact measurement was based on the

1. determination of the economic impact of HIV/AIDS morbidity and mortality in monetary terms (GH¢);
2. investigation of households' coping strategies; and
3. calculation of the direct and indirect costs of morbidity and mortality

Investigation of Households' Coping Strategies

Household coping strategies were analysed based on households' decision-making strategies. These decision-making were considered along dimensions such as consumption, health status, education, number of children, number of the aged, as well as the welfare of their extended family and unrelated community members. Households generally have resources with which they can pursue their welfare. These resources may include human capital (the number of household members, their education, and their earning capacity) and physical capital (savings, durable goods, productive assets, and land). They may use both the human and physical capital to generate income for making purchases subject to environmental constraints, which include the prices and quality of available goods and services such as food, housing, medical care and schooling.

The households were therefore specifically assisted by the research assistants to calculate their expenditure on household consumption of goods and services such as food, housing, medical care and schooling per month both before the on-set of illness or death and that after.

Computation of Direct and Indirect Costs of Morbidity and Mortality

The direct and indirect cost of an HIV/AIDS-related morbidity and mortality on a household was calculated using standard cost-benefit analysis. The direct costs of mortality was considered to include out-of-pocket medical care expenditure, travel expenses relating to medical care, the costs of funeral rites and other related expenditures.

The indirect costs of mortality was calculated from the foregone earnings of the deceased, whereas the foregone earnings of the diseased or incapacitated household member was calculated by working out the total number of lost work years by subtracting the age of the deceased or sick person from the average age of retirement, which was considered to be sixty for the Ghana Civil Service. For those with a regular income, annual income foregone was computed using, a five percent discount rate. The annual foregone earnings were multiplied by the number of lost work years to obtain the total foregone earnings. For those who had also held a supplementary job before their illness and death, these supplementary incomes were included in the calculations. Finally, in addition to the lost income of the deceased or incapacitated household member, the lost earnings of other household members who had to leave work to take care of the sick person was calculated in order to come up with the household's total foregone earnings.

Method of Data Analysis

Preliminary data analysis was done using the Statistical Package for the Social Sciences (SPSS®), version 20 while further data analyses was done using SAS®, version 9.4. Inductive and deductive reasoning were employed to arrive at explanations offered based on the trends of the analysis.

This study presents two levels of analysis (Preliminary Analysis and Further Analysis). The preliminary analysis comprises an extensive use of descriptive statistics and frequency distribution especially for the comparative differential impact of morbidity and mortality due to HIV/AIDS and 'Other illnesses/deaths'. The chapter on 'Further Analysis' comprises the modelling of the factors that determine the economic impact of HIV/AIDS-related morbidity and mortality on households, in particular, the impact on household income and changes in households' consumption levels, using regression analysis (Zhang et al., 2012). Economic variables were separately regressed with a number of socio-economic factors associated with the illness and/or death of an adult household member from HIV/AIDS or some other cause.

Variables used in the Analyses

Among the variables used in the analyses are "Type of Household", "Ethnicity of head of household", "Health Expenditure on children's education" and "Upkeep Expenditure on adults", among others. The rest of the variables used in the modeling process are presented in Appendices A1 to A3. Because of the problem of non-convergence of the models, some of the variables used in the modeling process had to be re-coded. The recoding was done such that the categories with the largest frequencies were maintained and those with smaller frequencies were merged into one. The non-convergence in the use of those variables could be traced to too many levels in them. Hence they had to be recoded to have less number of levels. The remaining variables presented in the preliminary analyses in Chapter Four but were absent in the models in Chapter Five are those whose inclusion did not make the models converge. Hence, they had to be entirely dropped to allow for convergence.

CHAPTER FOUR

RESULTS AND DISCUSSIONS

Introduction

In Chapter Three, several theoretical points that must always be taken into consideration when undertaking statistical modeling were presented. In this chapter, those points have been put into practice for the statistical modeling component of the work. The chapter is made up of preliminary analyses, which entail descriptive statistics, and further analyses, which also entail higher inferential analyses.

The second section examines the socio-demographic characteristics of respondents who are heads of households or their adult representatives. The socio-demographic characteristics cover variables such as age, highest educational level completed, marital status of heads of households, their occupation as at the time of data collection, religion, ethnicity, and type of household. The type of household refers to whether that household has experienced any recent illness or death due to HIV/AIDS or other illnesses in the past one year, or whether it has not experienced any illnesses and any disease in the past three months preceding the data collection. A household is considered to be an HIV/AIDS one if any member currently has HIV/AIDS or any member has recently (in the past one year) died of HIV/AIDS.

The third section takes a look at the socio-economic characteristics of surveyed households: which includes the household size, average monthly household income, ownership of assets, value of assets owned, and respondents' own perception of their economic status in their communities. The fourth section also examines the dependent population, i.e. children fifteen years old and below and adults above sixty years old within the household setting; their standard of living, upkeep, etc. The fifth section assesses the direct and indirect costs (in monetary terms) of illness and eventual death. The sixth section also assesses the socio-economic impacts of death on households, especially the dependent population and their coping strategies during illness and after death. The last section presents some summary data on the deceased in the households.

Socio-Demographic Characteristics of Respondents

The analysis of the socio-demographic characteristics of respondents was performed on comparative basis between households experiencing a recent HIV/AIDS illness or death, households experiencing illnesses and deaths due to causes other than HIV/AIDS, and those experiencing no illnesses and deaths. In view of that, a frequency table was constructed to cover sex, Table 3. Table 3 gives the distribution of socio-demographic characteristics of respondents. Key demographic variables captured are age, sex and marital status of respondents.

Whereas a little more of the males than females were interviewed from households experiencing illnesses and deaths due to other causes apart from HIV/AIDS and those with no recent experience of illness or death more females (i.e. 67.4%) happen to have been interviewed in the case of the HIV/AIDS sample. This might be a reflection of the general national HIV/AIDS distribution by gender.

The ages of the respondents were put into three age groups; i.e. 15-34, 35-49 and 50+, for the three categories of household under consideration (Table 4.1). Most of the respondents belonging to the households, “Other illnesses/deaths” and “No illness/ No death” were in the age group 15-34 years while most of the respondents in the HIV/AIDS households were within the age group 35-49 years.

Table 3: Sex, Age and Marital Status of Respondents

		Type of household		
		HIV/AIDS	Other Illnesses/deaths	No Illness/ No death
		n (%)	n (%)	n (%)
Sex				
	Female	122 (67.4)	92 (44.4)	101 (47.4)
	Male	59 (32.6)	115 (55.6)	112 (52.6)
	Total	181 (100)	207 (100)	213 (100)
Age of respondent				
	15-34	55 (30.4)	121 (58.5)	154 (72.3)
	35-49	97 (53.6)	61 (29.5)	47 (22.1)
	50+	29 (16.0)	25 (12.1)	12 (5.6)
	Total	181 (100)	207 (100)	213 (100)
Current Marital Status				
	Single	96 (53.0)	80 (38.6)	118 (55.4)
	Married	85 (47.0)	127 (61.4)	95 (44.6)
	Total	181 (100)	207 (100)	213 (100)

One striking feature about the study population, with respect to marital status, is that more than 50% of the respondents in “HIV/AIDS” households and “No illness/ No death” households are single. For those respondents who were married among the three categories of households, the highest proportion (61.4%) was recorded among respondents from “Other illnesses” households (Table 3).

The descriptive statistics of the age of respondents is presented in Table 4.1. The mean and median ages (34.14 and 34.00 respectively) of the total population of respondents are almost equal showing that the population of respondents is normally distributed. Whereas the minimum age of the “Other illnesses” and the “No illness/death” households was 18 years each, that of the HIV/AIDS household was 15.

Table 4: Descriptive Statistics on Age

Type of household	N	Median	Mean	Std. Deviation	Minimum	Maximum
HIV/AIDS	181	39.00	39.93	9.229	15	73
Other Illnesses/ deaths	207	31.00	33.45	10.655	18	65
No Illness/ No death	213	27.00	29.88	9.997	18	64
Total	601	34.00	34.14	10.796	15	73

The distribution of the relationship of respondents to the head of household is presented in Table 5. About half of the respondents interviewed from “HIV/AIDS” households were the heads of households whereas far less than half of their counterparts from the “Other illnesses/deaths” and “No Illnesses/deaths” households (39.6% and 32.9%) respectively were heads of households, (Table 5).

Another socio-demographic variable considered very important in the study, as far as the response variables are concerned, is ethnic groupings of respondents and is presented in Table 6.

Table 5: Relationship of Respondents to Heads of Household

	Type of household		
	HIV/AIDS	Other Illnesses/ deaths	No Illness/No death
	n (%)	n (%)	n (%)
Relationship to head of household			
Head	91 (50.3)	82 (39.6)	70 (23.9)
Spouse	40 (22.1)	52 (25.1)	43 (20.2)
Child/grand child	13 (7.2)	47 (22.7)	62 (29.1)
Other relation	25 (13.8)	24 (11.6)	34 (16.0)
Not related	12 (6.6)	2 (1.0)	4 (1.9)
Total	181 (100)	207 (100)	213 (100)

The largest represented ethnic group among respondents from "HIV/AIDS" and "Other illnesses/deaths" households was Akan (i.e. 38.1% from "HIV/AIDS", 38.6% from "Other illnesses"), Table 6. The Gas dominated among the "No illness/No death" household (34.3%). For those respondents who did not belong to any of the major ethnic groups stated in Table 6, majority of them were from "Other illnesses/deaths" households.

One other very important and desired characteristic of the study populations is the head of household's highest level of education completed. This is very important in helping to understand the underlying potentials of heads of households to overcome possible poor economic conditions that might be responsible for their health status. In view of this, respondents were asked to indicate their highest level of education completed, Table 6.

Table 6: Ethnicity and Level of Education of Respondents

	Type of household		
	HIV/AIDS	Other Illnesses/deaths	No Illness/No death
	n (%)	n (%)	n (%)
Ethnicity			
Akan	69 (38.1)	80 (38.6)	64 (30.0)
Ewe	44 (24.3)	39 (18.8)	51 (23.9)
Ga	42 (23.2)	49 (23.7)	73 (34.3)
Northerner	21 (11.6)	20 (9.7)	20 (9.4)
Other	5 (2.8)	19 (9.2)	5 (2.3)
Total	181 (100)	207 (100)	213 (100)
Highest level of education completed			
None	8 (4.4)	2 (1.0)	1 (0.5)
Primary	22 (12.2)	4 (1.9)	10 (4.7)
Middle/JHS/SHS	115 (63.5)	121 (58.5)	123 (57.7)
Voc./Comm.	19 (10.5)	19 (9.2)	28 (13.1)
Tertiary	13 (7.2)	58 (28.0)	50 (23.5)
Other	4 (2.2)	3 (1.4)	1 (0.5)
Total	181 (100)	207 (100)	213 (100)

The most represented educational level across the three categories of households was Middle/JHS/SHS. Within “Other illnesses” and “No illness/No death” households, respondents with Tertiary education followed those with Middle/JHS/SHS education in terms of proportion. The second largest group of respondents among “HIV/AIDS” households with respect to the highest level of education attained was those with Primary education. The type of household with the highest number of respondents with no education was the HIV/AIDS household. Quite a substantial proportion of respondents with vocational training were observed across the three groups of households.

One of the characteristics used in the selection of households for this study was a recent experience of death by the household and this is presented in Table 7.

Whereas there was death in 76.8% of HIV/AIDS households and almost 45.5% of these deaths were due to HIV/AIDS, there was death in 38.6% of the households with other illnesses and deaths. Respondents were asked if there had been any recent illness in their households in the past one year. Interestingly, about 45.9% of HIV/AIDS households experienced opportunistic illnesses and most of these illnesses, according to them, were HIV/AIDS-related.

By way of a follow-up, the respondents were asked if the opportunistic ailments their households had suffered were HIV/AIDS-related; and about 22.1% of respondents from HIV/AIDS households responded in the affirmative. With regards to their HIV status, whereas 36.7% of respondents from households with other illnesses were very sure they were HIV negative, probably because they went through Voluntary Counseling and Testing (VCT), as much as 62.8% of them were not sure of their HIV status. This number who know their status to be negative, appear to be in consonance with the rather low rate of VCT in this part of the world, Table 7.

Table 7: Morbidity and Mortality Characteristics of Households

	Type of household		
	HIV/AIDS	Other Illnesses/deaths	No Illness/No death
	n (%)	n (%)	n (%)
Whether any recent death in this household			
No	42 (23.2)	127 (61.4)	213 (100.0)
Yes	139 (76.8)	80 (38.6)	0 (0.0)
Total	181 (100)	207 (100)	213 (100)
If Yes whether HIV/AIDS-related			
Yes	30 (45.5)	2 (2.9)	0 (0.0)
No	36 (54.5)	66 (97.1)	0 (0.0)
Total	66 (100)	68 (100)	0 (0)
Whether any recent sickness or ailment in this household			
No	98 (54.1)	27 (13.0)	213 (100.0)
Yes	83 (45.9)	180 (87.0)	0 (0.0)
Total	181 (100)	207 (100)	213 (100)
If Yes whether it was HIV/AIDS-related			
Yes	40 (22.1)	2 (1.0)	0 (0.0)
No	43 (23.8)	178 (86.0)	0 (0.0)
Not Applicable	98 (54.1)	27 (13.0)	213 (100)
Total	181 (100)	207 (100)	213 (100)
Respondent's HIV status			
Positive	181 (100.0)	1 (0.5)	0 (0.0)
Negative	0 (0.0)	76 (36.7)	139 (65.3)
Don't know	0 (0.0)	130 (62.8)	74 (34.7)
Total	181 (100)	207 (100)	213 (100)

Socio-economic Characteristics of Surveyed Households

Presented in Table 8 are the measures of central tendency and dispersion for the variable “Household size”.

Table 8: Descriptive Statistics on Household Size

Type of household		Household Size
HIV/AIDS		
	N	181
	Median	4
	Mean	4.74
	Std. Deviation	2.589
	Minimum	1
	Maximum	15
Other Illnesses/deaths		
	N	207
	Median	4
	Mean	4.85
	Std. Deviation	2.077
	Minimum	1
	Maximum	14
No Illness/No death		
	N	213
	Median	4
	Mean	4.16
	Std. Deviation	1.594
	Minimum	1
	Maximum	11
Total		
	N	601
	Median	4
	Mean	4.57
	Std. Deviation	2.118
	Minimum	1
	Maximum	15

The HIV/AIDS households had a mean size of 4.7 with a mode of 4 whereas the households with other illnesses had a mean of 4.9 and a mode of 4. "No illness/death" households had a mean of 4.2 and a mode of 4 as well, (Table 8).

In Table 9, Respondents were asked to provide their household sizes.

These are also meant to help the researcher better explore factors which have contributory effects on the impacts households face due to illnesses and/or deaths. Due to varying household sizes, the figures were grouped with intervals of five each.

Table 9: Household Size and Asset Ownership of Surveyed Households

	Type of household		
	HIV/AIDS	Other Illnesses/deaths	No Illness/No death
	n (%)	n (%)	n (%)
Household size			
1 to 5	126 (69.9)	142 (68.6)	180 (84.5)
6 to 10	49 (27.1)	60 (29.0)	32 (15.0)
15 or more	6 (3.3)	5 (2.4)	1 (0.5)
Subtotal	181 (100)	207 (100)	213 (100)
Total value of assets owned (in GH¢)			
< 100	124 (68.5)	15 (7.2)	70 (32.9)
100 – 999	14 (7.7)	76 (36.7)	40 (18.8)
1,000 - 9,999	31 (17.1)	51 (24.6)	54 (25.4)
10,000 - 99,999	11 (6.1)	53 (25.6)	44 (20.7)
100,000 and above	1 (0.6)	12 (5.8)	5 (2.3)
Subtotal	181 (100)	207 (100)	213 (100)

A great majority of the three types of household under investigation (i.e. 69.6% of the "HIV/AIDS" households, 68.5% of the "Other illnesses" households and about 84.5% of the "No illnesses/diseases" households) had a size of 1-5, Table 9. Some small proportions of all the three categories of household had very large sizes (i.e.15 or more).

One of the key variables that were suspected to be responsible for determining the intensity of the impact that households faced due to illness and/or death is household income (Zhang et al., 2012). It is also the determining factor for whether any orphaned children have to leave school to fend for themselves or not. One of the questions in the questionnaire therefore sought the average household income from respondents, Table 9.

Table 10: Socio-economic Characteristics of Surveyed Households

	Type of household		
	HIV/AIDS n (%)	Other Illnesses/deaths n (%)	No Illness/No death n (%)
Household owns a House			
No	132 (72.9)	93 (44.9)	108 (50.7)
Yes	49 (27.1)	114 (55.1)	105 (49.3)
Subtotal	181 (100)	207 (100)	213 (100)
Household owns a Farm Land			
No	166 (91.7)	165 (79.7)	179 (84.0)
Yes	15 (8.3)	42 (20.3)	34 (16.0)
Subtotal	181 (100)	207 (100)	213 (100)
Household owns a Building Land			
No	153 (84.5)	145 (70.0)	159 (74.6)
Yes	28 (15.5)	62 (30.0)	54 (25.4)
Subtotal	181 (100)	207 (100)	213 (100)
Household owns a Car			
No	161 (89.0)	123 (59.4)	137 (64.3)
Yes	20 (11.0)	84 (40.6)	76 (35.7)
Subtotal	181 (100)	207 (100)	213 (100)
Household owns Livestock			
No	139 (76.8)	85 (41.1)	118 (55.4)
Yes	42 (23.2)	122 (58.9)	95 (44.6)
Subtotal	181 (100)	207 (100)	213 (100)

Asset ownership is another of the factors that this researcher believes would help respondent households severely affected by illness and/or death to mitigate the hardship due to lost income from the illness and/or death.

For a preliminary look at the effects of this factor on the households, respondents were asked if their households owned a house. Whereas only 27.1% of “HIV/AIDS” households owned at least one house, majority of the “Other illnesses” household (55.1%) and almost half (49.3%) of the “No illness/death” households owned at least one house, Table 10. This trend reflects in the ownership of other assets. For instance, only 8.3% of the “HIV/AIDS” households, compared to about 20.3% of the “Other illnesses” and 16% of the “No illness/death” households, owned a farm land.

Similarly in Table 10, whereas only about 15.5% of the “HIV/AIDS” households owned a building land, about 30% of the “Other illnesses” households and about 25.4% of the “No Illness/death” households owned a building land. Also, whereas only 11% of the “HIV/AIDS” households owned a car, about 40.6% and 35.7% of the “Other illnesses” and “No illness/death” households respectively owned a car.

In Table 11, respondents were asked to indicate the total value of assets owned and measures of central tendency and dispersion is presented on it. The average value of assets owned by other illnesses/deaths households was the largest among the three households considered. However, the largest maximum value of total assets owned was from households with no illness/death.

Table 11: Descriptive Statistics on Total Value of Assets Owned (in GH¢)

Type of household	N	Median	Mean	Std. Deviation	Minimum	Maximum
HIV/AIDS	181	0.00	2,415.72	9,187.36	0.00	100,000.00
Other Illnesses/deaths	207	2,000.00	20,262.00	53,681.35	0.00	500,000.00
No Illness/No death	213	500.00	8,460.65	19,484.88	0.00	120,000.00
Total	601	400.00	10,704.82	34,682.50	0.00	500,000.00

Characteristics of the Dependent Population

An important characteristic in this study is the dependent population, i.e. household members 15 years old or less and those above 60 years. They are considered to be the ones most hardly hit by any shocks in the households due to illnesses or deaths since they mostly depend on the working age group for their livelihood. In view of this, respondents were asked if there were any children age 15 years or less in their respective households and this is presented in Table 12.

Table 12: The Dependent Population (Children under 15)

	Type of household		
	HIV/AIDS	Other Illnesses/deaths	No Illness/No death
	n (%)	n (%)	n (%)
Any children less than 15 years in the household?			
Yes	145 (80.1)	140 (67.6)	90 (42.3)
No	36 (19.9)	67 (32.4)	123 (57.7)
Total	181 (100)	207 (100)	213 (100)

The total expenditure incurred on children is made up of expenditure incurred on their education, health and upkeep and this is presented in Table 13.

Table 13: Expenditure (in GH¢) Incurred on Children after Disease Onset

	Type of household		
	HIV/AIDS	Other Illnesses/deaths	No Illness/No death
	n (%)	n (%)	n (%)
Expenditure on children's education after the onset of disease/death			
None	35 (19.3)	67 (32.4)	123 (57.7)
<100	92 (50.8)	49 (23.7)	89 (41.8)
100 – 999	48 (26.5)	73 (35.3)	1 (0.5)
1,000 and above	6 (3.3)	18 (8.7)	0 (0.0)
Total	181 (100)	207 (100)	213 (100)
Expenditure on children's health after the onset of disease/death			
None	35 (19.3)	67 (32.4)	123 (57.7)
<100	99 (54.7)	52 (25.1)	80 (37.6)
100 and above	47 (26.0)	88 (42.5)	10 (4.7)
Total	181 (100)	207 (100)	213 (100)
Expenditure on children's upkeep after the onset of disease/death			
None	35 (19.3)	67 (32.4)	123 (57.7)
<100	6 (3.3)	0 (0.0)	0 (0.0)
100 – 999	117 (64.6)	29 (14.0)	70 (32.9)
1,000 and above	23 (12.7)	111 (53.6)	20 (9.4)
Total	181 (100)	207 (100)	213 (100)
Total expenditure on children after the onset of disease/death			
None	35 (19.3)	67 (32.4)	123 (57.7)
<100	0 (0.0)	0 (0.0)	0 (0.0)
100 – 999	97 (53.6)	26 (12.6)	70 (32.9)
1,000 and above	49 (27.1)	114 (55.1)	20 (9.4)
Total	181 (100)	207 (100)	213 (100)

Table 14: Descriptive Statistics on Expenditure (in GH¢) Incurred on Children after Disease Onset

	Type of household			Total
	HIV/AIDS	Other Illnesses/ deaths	No Illness/ No death	
Expenditure on children's education after disease onset				
N	181	207	213	601
Median	30.00	50.00	0.00	25.00
Mean	178.02	253.41	13.44	145.66
Std. Deviation	349.79	404.30	20.31	321.72
Minimum	0.00	0.00	0.00	0.00
Maximum	2240.00	2550.00	200.00	2550.00
Expenditure on children's health after the onset of disease/death				
N	181	207	213	601
Median	50.00	60.00	0.00	30.00
Mean	110.45	115.59	21.14	80.57
Std. Deviation	174.05	165.16	31.59	144.06
Minimum	0.00	0.00	0.00	0.00
Maximum	800.00	1200.00	200.00	1200.00
Expenditure on children's upkeep after the onset of disease/death				
N	181	207	213	601
Median	300.00	1000.00	0.00	250.00
Mean	452.88	1944.73	256.53	897.12
Std. Deviation	539.69	2570.24	479.66	1737.99
Minimum	0.00	0.00	0.00	0.00
Maximum	2500.00	11100.00	2100.00	11100.00
Total expenditure on children after the onset of disease/death				
N	181	207	213	601
Median	485.00	1180.00	0.00	445.00
Mean	741.35	2313.73	291.10	1123.35
Std. Deviation	806.75	2920.69	499.85	1997.77
Minimum	0.00	0.00	0.00	0.00
Maximum	4840.00	12600.00	2139.00	12600.00

After the onset of diseases/deaths, about 50% of HIV/AIDS and 41.8% of no illness/death households spent less than GH¢ 100. More than one-third of the total proportion of “Other illnesses/deaths” households spent between GH¢ 100 and GH¢ 1000. The trend was the same with the expenditure made by the different groups of households on the health of their children, Table 13.

The questions asked concerning the expenditure made by households on their children’s upkeep showed that 64.6% of “HIV/AIDS” households spent between GH¢ 100 and GH¢ 1,000 and also 32.9% of “No illness/death” made the same expenditure. Households made up of “Other illnesses/deaths” had 53.6% of their total proportion spending GH¢ 1,000 and more on the upkeep of their children after the onset of diseases/death, Table 13.

In Table 15 is the expenditure incurred on adults’ upkeep and death after the onset of illness. Among the three groups of households, the households with the highest and lowest average expenditure on their children’s education, health, up keep and total expenditure after the onset of disease/death were “Other illnesses/death” households and “No illness/no death” households respectively.

In all the three categories of households, quite a considerable proportion of households spent between GH¢100 to GH¢1,000 on adult health after the onset of disease/death. Expenditure on adult upkeep after the onset of disease/death was between GH¢100 to GH¢1,000 among a significant number of “HIV/AIDS” households (47) and “No illness/no death” households (28), while a substantial number of “Other illnesses/deaths” households (28) spent GH¢1,000 and more, Table 15.

The total expenditure on adults after the onset of disease/death increased by 22.1%, 10.1% and 13.1% for “HIV/AIDS”, “Other illnesses/deaths” and “No illness/no death” households respectively, and those increases ranged between GH¢100 to GH¢1,000, Table 15.

Table 15: Expenditure Incurred on Adults after Disease Onset (GH¢)

	Type of household		
	HIV/AIDS	Other Illnesses/deaths	No Illness/No death
	n (%)	n (%)	n (%)
Expenditure on adult's health after the onset of disease/death			
None	103 (56.9)	135 (65.2)	170 (79.8)
<100	10 (5.5)	11 (5.3)	6 (2.8)
100 – 999	62 (34.3)	50 (24.2)	36 (16.9)
1,000 and above	6 (3.3)	11 (5.3)	1 (0.5)
Total	181 (100)	207 (100)	213 (100)
Expenditure on adult's upkeep after the onset of disease/death			
None	103 (56.9)	135 (65.2)	170 (79.8)
<100	0 (0.0)	0 (0.0)	0 (0.0)
100 – 999	47 (26.0)	23 (11.1)	28 (13.1)
1,000 and above	31 (17.1)	49 (23.7)	15 (7.0)
Total	181 (100)	207 (100)	213 (100)
Total expenditure on adult after the onset of disease/death			
None	103 (56.9)	135 (65.2)	170 (79.8)
<100	0 (0.0)	0 (0.0)	0 (0.0)
100 – 999	40 (22.1)	21 (10.1)	28 (13.1)
1,000 and above	38 (21.0)	51 (24.6)	15 (7.0)
Total	181 (100)	207 (100)	213 (100)

The group with the highest number and proportion of households that had an expenditure of GH¢1,000 or more on adults after the onset of disease was "Other illnesses/death" households, Table 15. It can be seen that expenditure was incurred on health in households where there was no illness/death. This expenditure was incurred on preventive health, such as medical check-ups.

Table 16: Descriptive Statistics on Expenditure on Adults after Disease Onset (in GH¢)

	Type of household			
	HIV/AIDS	Other Illnesses/ deaths	No Illness/ No death	Total
Health expenditure on adults after disease onset				
N	181	207	213	601
Median	0	0	0	0
Mean	170.58	228.12	53	148.73
Std. Deviation	399.87	758.74	146.5	508.7
Minimum	0	0	0	0
Maximum	3700	5500	1100	5500
Upkeep expenditure on adults after disease onset				
N	181	207	213	601
Median	0	0	0	0
Mean	424.36	982.85	227.65	547
Std. Deviation	791.5	2727.29	789.66	1751.66
Minimum	0	0	0	0
Maximum	4000	30000	5600	30000
Total expenditure on adults after disease onset				
N	181	207	213	601
Median	0	0	0	0
Mean	594.94	1210.97	280.66	695.73
Std. Deviation	1013.86	3203.12	886.59	2065.07
Minimum	0	0	0	0
Maximum	5100	34500	6120	34500

Presented in Table 15 are the measures of central tendency and dispersion on expenditure incurred on adults' upkeep and death after the onset of illness.

The mean expenditure on adult health after disease onset was highest in "Other illnesses/deaths" households (GH¢228.12), followed by "HIV/AIDS" households (GH¢170.58). "No illness/No death" households presented the lowest average expenditure (GH¢53.00) among the three households, Table 16.

"Other illnesses/deaths" households recorded the highest average monthly expenditure on upkeep of adults after onset of disease, with "HIV/AIDS" households following. "No illness/no death" households recorded the lowest mean expenditure among the three households (GH¢227.65), Table 16.

Overall, "Other illnesses/deaths" households recorded the highest average expenditure (GH¢1210.97) on adults after onset of disease, followed by "HIV/AIDS" households (GH¢594.94). "No illness/no death" households spent the least income (GH¢280.66) on adults after onset of disease. Health expenditure incurred in households where there was no illness/deaths was actually incurred on medical check-ups.

Other Related Variables Studied

There were other related variables which were included in the study from the beginning. However because they were discovered not to have had any significant effect on the models, particularly because they were discovered to be hindering the models from converging, they were dropped completely. These variables include religion, residential status, occupation, number of children and adults in the household and households' coping strategies during illness and after death.

Summary of Preliminary Analysis

It is found that expenditure on health care or diseases/deaths, other than HIV, is the highest among the three categories of household. This indicates that it is more expensive dealing with non-HIV/AIDS-related illnesses. This is followed by HIV/AIDS illnesses.

The HIV/AIDS households incurred more expenditure on their dependent children's health, in the form of check-ups, than their education and upkeep, contrary to what pertains in the other two categories of household. They also incurred far less total expenditure on their dependent children than the other two categories of household before the onset of the disease but not after. Then HIV/AIDS category of households also incurred a lower cost of medical treatment, monthly income loss and funeral expenses than their "Other Illness/Deaths" category of households. However, they incurred a higher travel cost than their counterparts from the "Other Illness/Deaths" category of households.

Very few of the HIV/AIDS category of households owned assets such as house, farm land, building land, a car, livestock, etc., compared to their counterparts in the other two categories of household. On the total value of assets owned, the HIV/AIDS category of households had the least compared to the other two categories. Most of them had dependents than the other two category of households.

Among the highlights of the chapter was the fact that HIV/AIDS households were more predominantly headed by females than the other two categories. Although the heads of the HIV/AIDS category of households are a shade older than their counterparts from the other two categories of household, they were the only category of households that had the youngest heads. The HIV/AIDS households incurred more deaths than the "Other Illnesses/Deaths" households.

Further Analysis

In this section, we obtain a generalized linear mixed model of dependents' likelihood of reallocating time, dependents working harder to substitute for lost household income, dependents leaving job to care for the sick on socio-demographic characteristics in the event of illness/death. An additional generalized linear mixed model was run for economic determinants of households' likelihood of reducing expenditure on household size, total value of assets, expenditure on children's education as well as expenditure on adults' health.

We make use of "PROC GLIMMIX" procedure in SAS® in this chapter. In each of the five models, an assessment was made between the fit of MMPL and RMPL as well as the interpretation of the models as far as the fixed and random effects were concerned.

The independent variables listed above were finally selected from among a larger list of independent variables as it turned out that they were the only independent variables which brought about convergence during the modeling process.

Ascertaining the Socio-Demographic Determinants of the Impacts of Illnesses and Deaths on Households

One of the objectives of this study is to ascertain the extent of impact of illness and/or death on the dependent population within the household setting. Whenever there is illness and/or death, the entire household is expected to be affected one way or the other, (Cohen, 1993; HSRC, 2001a; Rugalema, 1999a). However, the dependent population is expected to be more affected, since they depend almost entirely on the working population within the household setting for their livelihood. This dependent population is defined as persons outside the economically active age group. This age group includes children who are less than fifteen years as well as retired persons above the age of sixty.

Dependents' likelihood of reallocating time

The general form of the Generalized Linear Mixed Model (GLMM) is given as

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{y}$, which is the outcome vector, is an $N \times 1$ column vector (with an underlying logistic link function of $(.) = log_e\left(\frac{p}{1-p}\right)$,

$\boldsymbol{X}$ is an $N \times p$ matrix of the p predictor variables,

$\boldsymbol{\beta}$ is a $p \times 1$ vector of fixed-effects regression coefficients,

$\boldsymbol{Z}$ is the $N \times q$ design matrix for the q random effects (the random compliment to the fixed $\boldsymbol{X}$),

$\boldsymbol{\gamma}$ is a $q \times 1$ vector of the random effects (the random complement to the fixed $\boldsymbol{\beta}$), and

$\boldsymbol{\varepsilon}$ is an $N \times 1$ column vector of the residuals, that part of y that is not explained by the model $\boldsymbol{X\beta} + \boldsymbol{Z\gamma}$. That is,

$$\underbrace{\boldsymbol{y}}_{N\times1} = \overbrace{\underbrace{\boldsymbol{X}}_{N\times p}\quad\underbrace{\boldsymbol{\beta}}_{p\times1}}^{N\times1} + \overbrace{\underbrace{\boldsymbol{Z}}_{N\times q}\quad\underbrace{\boldsymbol{\gamma}}_{q\times1}}^{N\times1} + \underbrace{\boldsymbol{\varepsilon}}_{N\times1}$$

There were 601 participants in the study. Thus $n = 601$. There are $p = 13$ parameters to be estimated. The parameters are the coefficients of various levels of five independent variables.

Sex (1=male, 0=female), Age (2=15-34, 1=35-49, 0=50+), Marital Status (1=single, 0=married), Education (2=Up to JSS, 1=SHS, 0=Post Sec and others), Any recent HIV-related illness in household (1=No, 0=Yes) and intercept component.

The condition of the respondents at the household level is one of HIV/AIDS, Other Illnesses/Deaths and No Illness/No Death). All these three are mutually exclusive. Thus, $q = 3$.

Hence the design matrix $\boldsymbol{X}$ which is 601×13, is given as

$$\boldsymbol{X} = \begin{array}{c} \begin{array}{cccccc} \text{Intercept} & \text{Sex} & \text{Age} & \text{Mar. Stat.} & \text{Educ.} & \text{Oth. Illn.} \end{array} \\ \begin{pmatrix} 1_1 & 1\ \ 0 & 1\ \ 0\ \ 0 & 0\ \ 1 & 0\ \ 1\ \ 0 & 0\ \ 1 \\ 1_2 & 1\ \ 0 & 1\ \ 0\ \ 0 & 0\ \ 1 & 0\ \ 1\ \ 0 & 0\ \ 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1_{44} & 0\ \ 1 & 1\ \ 0\ \ 0 & 1\ \ 0 & 1\ \ 0\ \ 0 & 1\ \ 0 \\ 1_{45} & 1\ \ 0 & 0\ \ 1\ \ 0 & 1\ \ 0 & 0\ \ 0\ \ 1 & 0\ \ 1 \\ 1_{46} & 0\ \ 1 & 0\ \ 0\ \ 1 & 1\ \ 0 & 0\ \ 0\ \ 1 & 1\ \ 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1_{601} & 0\ \ 1 & 0\ \ 1\ \ 0 & 0\ \ 1 & 0\ \ 1\ \ 0 & 1\ \ 0 \end{pmatrix} \end{array}$$

And the design matrix $\boldsymbol{Z}$ is given as

$$\boldsymbol{Z}=\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1_{44} & 0 & 0 \\ 0_{45} & 1 & 0 \\ 0_{46} & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1_{601} & 0 & 0 \end{pmatrix}_{601\times 3} \quad \text{and the vector } \boldsymbol{y} \text{ is given as } \boldsymbol{y}=\begin{pmatrix} 0_1 \\ 0_2 \\ \vdots \\ 1_{44} \\ 1_{45} \\ 1_{46} \\ \vdots \\ 0_{601} \end{pmatrix},$$

where $1=$ "No re-allocation of time" and $0=$ "Re-allocation of time" for the MMPL model. It must be noted that the vector, $\boldsymbol{y}$, is not a zero vector. Thus

$$\begin{pmatrix} 0_1 \\ 0_2 \\ \vdots \\ 1_{44} \\ 1_{45} \\ 1_{46} \\ \vdots \\ 0_{601} \end{pmatrix}=$$

$$\begin{array}{c} \begin{array}{cccccc} \text{Intercept} & \text{Sex} & \text{Age} & \text{Mar. Stat.} & \text{Educ.} & \text{Oth. Illn.} \end{array} \\ \begin{pmatrix} 1_1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1_2 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ \vdots & \vdots & & & \vdots & & & \vdots & & \vdots & & & \vdots \\ 1_{44} & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1_{45} & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1_{46} & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ \vdots & \vdots & & & \vdots & & & \vdots & & \vdots & & & \vdots \\ 1_{601} & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} \end{array} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{13} \end{pmatrix}+$$

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1_{44} & 0 & 0 \\ 0_{45} & 1 & 0 \\ 0_{46} & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1_{601} & 0 & 0 \end{pmatrix}_{601\times 3} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}+\begin{pmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \\ \vdots \\ \boldsymbol{\varepsilon_{601}} \end{pmatrix}$$

The first, second and last responses in the database are zeros. But between the second response and the past-but-one, there are non-zeroes.

Solving the above equation gives the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ vector solutions as

$$\boldsymbol{\beta} = \begin{pmatrix} 2.3291 \\ 0.3572 \\ 0.0000 \\ 0.2009 \\ -0.6476 \\ 0.0000 \\ -0.2873 \\ 0.0000 \\ 0.7033 \\ 0.5488 \\ 0.0000 \\ -0.8181 \\ 0.0000 \end{pmatrix} \text{ and } \boldsymbol{\gamma} = \begin{pmatrix} -2.0377 \\ 0.9183 \\ 1.1194 \end{pmatrix}$$

for the MMPL model and

$$\boldsymbol{\beta} = \begin{pmatrix} 2.3268 \\ 0.3623 \\ 0.0000 \\ 0.1905 \\ -0.6475 \\ 0.0000 \\ -0.2860 \\ 0.0000 \\ 0.7038 \\ 0.5431 \\ 0.0000 \\ -0.8138 \\ 0.0000 \end{pmatrix} \text{ and } \boldsymbol{\gamma} = \begin{pmatrix} -2.0531 \\ 0.9271 \\ 1.1260 \end{pmatrix}$$

for the RMPL model. For more clarity of the representation of these estimates, $\boldsymbol{\beta}$ is presented in Table 17 with respective p-values.

As discussed in Chapter Three, $\boldsymbol{\gamma} \sim N(\mathbf{0}, \boldsymbol{G})$ and $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$, where $\boldsymbol{G}$ and $\boldsymbol{R}$ are the G–side (i.e. Gamma-side) variance-covariance matrix and R-side (Residual-side) variance-covariance matrix and $\mathbf{0}$ is a zero matrix.

The matrices $\boldsymbol{G}$ and $\boldsymbol{R}$ are identified, for the MMPL model, as

$$\boldsymbol{G} = \mathbf{I}_3\sigma_\gamma^{\mathbf{2}} = \begin{pmatrix} 2.0377 & 0 & 0 \\ 0 & 0.9183 & 0 \\ 0 & 0 & 1.1194 \end{pmatrix}$$

and

$$\boldsymbol{R} = \mathbf{I}_3\sigma_\varepsilon^{\mathbf{2}} = \begin{pmatrix} 2.1252 & 0 & 0 \\ 0 & 0.035 & 0 \\ 0 & 0 & 0.7229 \end{pmatrix}$$ for the MMPL model,

while

$$\boldsymbol{G} = \mathbf{I}_3\sigma_\gamma^{\mathbf{2}} = \begin{pmatrix} 2.0531 & 0 & 0 \\ 0 & 0.9271 & 0 \\ 0 & 0 & 1.1260 \end{pmatrix}$$ and

$$\boldsymbol{R} = \mathbf{I}_3\sigma_\varepsilon^{\mathbf{2}} = \begin{pmatrix} 3.2372 & 0 & 0 \\ 0 & 0.0336 & 0 \\ 0 & 0 & 0.7306 \end{pmatrix}$$ for the RMPL model.

The significance of $\boldsymbol{G}$ and $\boldsymbol{R}$ is presented in Tables 18 and 19 for both the MMPL and RMPL models.

Table 17: Socio-demographic Determinants of Dependents' Likelihood of Reallocating Time (Model 1)

Predictor Variables	Estimation Approach							
	Maximum Pseudo-Likelihood				Residual Pseudo-Likelihood			
	B	S.E.	Test Statistic	P	B	S.E.	Test Statistic	p
Intercept	2.329	1.24	1.88	0.201	2.327	1.385	1.68	0.235
Sex								
Female	0.357	0.257	1.39	0.165	0.362	0.258	1.4	0.161
Male	0	.	.	.	0	.	.	.
Age								
15-34	0.201	0.454	0.44	0.658	0.191	0.456	0.42	0.677
35-49	-0.65	0.426	-1.52	0.129	-0.648	0.429	-1.51	0.131
50+	0	.	.	.	0	.	.	.
Marital Status								
Single	-0.29	0.283	-1.02	0.31	-0.286	0.284	-1.01	0.315
Married	0	.	.	.	0	.	.	.
Education								
Up to JSS	0.703	0.832	0.85	0.398	0.704	0.836	0.84	0.4
SHS	0.549	0.829	0.66	0.508	0.543	0.833	0.65	0.515
Post Sec and Others	0	.	.	.	0	.	.	.
Any recent illness in household								
No	-0.82	0.34	-2.41	0.016	-0.814	0.342	-2.38	0.018
Yes	0	.	.	.	0	.	.	.

Table 18: Covariance Parameter Estimates for Model 1

Covariance Parameter	Estimation Approach					
	Maximum Pseudo-Likelihood			Pseudo Pseudo-Likelihood		
	Subject	Est	S E	Subject	Est	S E
TYPE		2.1252	1.7817		3.2372	3.2948
AR(1)	Intercept	0.035	0.0419	Intercept	0.0336	0.0419
Residual		0.7229	0.0419		0.7306	0.0425

PROC GLIMMIX identifies the variable 'Reallocation of time' as response variable and binary in nature. The estimation technique specified in the models are maximum marginal pseudo-likelihood (METHOD=MMPL) and residual marginal pseudo-likelihood (METHOD=RMPL) with a subject-specific expansion, respectively, as specified in Chapter Three.

Table 19: Solutions for Random Effects for Model 1

Type of Household	Estimation Approach							
	Maximum Pseudo-Likelihood				Pseudo Pseudo-Likelihood			
	Estimate	S E Pred	t Value	Pr > |t|	Estimate	S E Pred	t Value	Pr > |t|
HIV/AIDS	2.04	0.86	-2.36	0.02	2.05	1.06	-1.94	0.05
Other Illnesses/ Deaths	0.92	0.87	1.05	0.29	0.3	1.06	0.87	0.38
No Illness/ No Death	1.12	0.87	1.29	0.2	1.13	1.06	1.06	0.29

They both have logit link functions (Appendices C1.1 and C2.1). The "Class Level Information" table, for both models, lists the levels of the variables specified in the CLASS statement and the ordering of the levels (Appendices C1.2 and C2.2). From there, there are six variables listed, one response and the remaining five explanatory variables, respectively. Three of the variables have three levels whereas the remaining four had two levels respectively. From the "Number of Observations" table in the SAS output (Appendices C1.3 and C2.3), the number of observations read and used in the analysis is 601, respectively. In the “Dimensions” table are listed the size of related matrices (Appendices C1.5 and C2.5). The X-matrix contains 13 columns, one of which is an intercept, and the remaining 12 represent the levels of the fixed effects variables all together. The random effect is made up of G-side and R-side covariance parameters of dimension 1 and 2, respectively, for each of the models.

The “Optimization Information” table (Appendices C1.6 and C2.6) in the SAS output presents information about the methods and size of the optimization problem. The maximum number of observations utilized per subject is 601, implying that every information was utilized by every subject in obtaining the parameters of the model. The optimization technique for both the MMPL and RMPL forms of the GLMM with binary data is the Newton-Raphson with Ridging. The “Iteration History” table (Appendices 1.7 and 2.7) also in the SAS output displays information about the progress of the optimization process. After the initial optimization, the GLIMMIX procedure performs 18 updates before the convergence criterion is met for each of the two models. At convergence, the largest absolute value of the gradient is almost zero, indicating the fact that the process stops at an extremum of the objective function for each of the models.

The “Model Fit Statistics” component which is presented in Appendices C1.8 and C2.8 and Table 20, gives information about the fitted model. The -2Log

Likelihood in the final MMPL model is 3084.57 while the -2 Residual Log Pseudo-Likelihood of the RMPL model is 3081.66. The ratio of the generalized chi-square statistic and its degree of freedom is approximately 1 for both models and this is a measure of the maximum variability in the marginal distribution of the underlying data. This implies that overdispersion is absent in the model.

The "Covariance Parameter Estimates" table displays estimates and asymptotic estimated standard errors for all covariance parameters for both models (Appendix 5.4) and the variance-covariance matrix of the MMPL model is presented above. The random effect, TYPE, representing type of household, is estimated at 2.1252 with a standard error of 1.7817 for the MMPL model and 3.2372 with a standard error of 3.2948 for the RMPL model (Appendices C1.9 and C2.9).

Table 20: Model Fit Statistics for Model 1

Model Fit Statistics	Estimation Approach	
	Maximum Pseudo-Likelihood	Residual Pseudo-Likelihood
-2 log	3081.7	3084.5
Pseudo-AIC	3103.7	3106.6
Pseudo-AICC	3104.1	3107.1
Pseudo-BIC	3096.7	3151.9
Pseudo-CAIC	3107.7	3162.9
Pseudo-HQIC	3086.6	3122.4

This simply implies that, all things being equal, the household effect, also representing the disease effect, is not only higher in the RMPL model, but the specific household means also vary from the population mean more drastically than is the case for the MMPL model. The RMPL with a higher standard error represents a greater deviation from the mean intercept and slope than the MMPL model.

In the tables in Appendices C1.10 and C2.10, and also in Table 17, the estimates of the fixed effects, their standard errors and their *p*-values are presented. Of all the explanatory variables utilized among the fixed effects, only “Recent illness in household” significantly ($p<0.05$) explains dependents’ likelihood of reallocating time, for both models.

The statistically significant variance components of the model for HIV/AIDS category of households (-2.038 for the MMPL model and -2.053 for RMPL model in Appendices C1.12 and C2.12 respectively and Table 19) shows that there is the unlikelihood of reallocating time as a result of illness or death and this varies across HIV/AIDS households ($p<0.05$) at 5% level of significance for the MMPL model in Appendix 5.1 and it does not vary across HIV/AIDS households ($p>0.05$) for the RMPL model. However, the variability in the likelihood of reallocating time across the other two categories of households (“Other Illnesses/Deaths” and “No Recent Illness/Deaths”) is not significant.

Given that from Table 19, $\gamma_1 = '\text{HIV/AIDS}$ households’ is the only significant component of the $\boldsymbol{\gamma}$ vector ($p<0.05$), the entire model is interpreted such that the prediction of the likelihood of occurrence of the response variables by the fixed effects is applicable to only ‘HIV/AIDS households’. However, for the other category of households, the random effect does not apply due to the fact that the random effect is not statistically significant for them.

Thus for households where there is no recent HIV-related illness or death, there is a less likelihood that a dependent will reallocate time. All the fixed effects variables are not significant determinants of dependents' likelihood of reallocating time, except 'recent illness in the household' ($p<0.05$ for both MMPL and RMPL models). This can be seen in Appendices C1.10 and C2.10, and also in Table 19. Thus for both models, the unlikelihood of reallocating time, due to recent illness in households, varies for HIV/AIDS households, implying that some households' dependents are more unlikely to reallocate time than others, for both models.

Dependents Working Harder to Substitute for Lost Household Income

For this variable, there are $p = 7$ parameters to be estimated. The parameters are the coefficients of the three relevant independent variables. These are Sex (1=male, 0=female), Marital Status (1=single, 0=married), Ethnicity (1=Akan, 0=Non-Akan).

The condition of the respondents at household level is one of HIV/AIDS, Other Illnesses/Deaths and No Illness/No Death. Thus, $q = 3$. All these are mutually exclusive. Hence the design matrix $\boldsymbol{X}$ is given as

$$
\boldsymbol{X} = \begin{array}{c}
\begin{array}{ccccccc} \text{Intercept} & \text{Sex} & & \text{Marit. Stat.} & & \text{Ethnicity} & \end{array} \\
\begin{pmatrix}
1_1 & 1 & 0 & 0 & 1 & 0 & 1 \\
1_2 & 1 & 0 & 0 & 1 & 0 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{44} & 0 & 1 & 1 & 0 & 1 & 0 \\
1_{45} & 1 & 0 & 0 & 1 & 0 & 1 \\
1_{46} & 0 & 1 & 1 & 0 & 0 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{601} & 0 & 1 & 0 & 1 & 0 & 1
\end{pmatrix}_{601\times 7}
\end{array}
$$

The design matrix $\boldsymbol{Z}$ and response vector $\boldsymbol{y}$ are similarly respectively given as

$$
\boldsymbol{Z} = \begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{44} & 0 & 0 \\
0_{45} & 1 & 0 \\
0_{46} & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{601} & 0 & 0
\end{pmatrix}_{601\times 3} \quad \text{and } \boldsymbol{y} = \begin{pmatrix}
0_1 \\ 0_2 \\ \vdots \\ 1_{44} \\ 1_{45} \\ 0_{46} \\ \vdots \\ 0_{601}
\end{pmatrix},
$$

where $1 =$ 'Works harder to substitute for lost income' and $0 =$ 'Does not work harder to substitute for lost income', for the MMPL model. It should be noted that the vector y is not a zero vector. Thus

$$
\begin{pmatrix}
0_1 \\ 0_2 \\ \vdots \\ 1_{44} \\ 1_{45} \\ 0_{46} \\ \vdots \\ 0_{601}
\end{pmatrix} = \begin{array}{c}
\begin{array}{ccccccc} \text{Intercept} & \text{Sex} & & \text{Marit. Stat.} & & \text{Ethnicity} & \end{array} \\
\begin{pmatrix}
1_1 & 1 & 0 & 0 & 1 & 0 & 1 \\
1_2 & 1 & 0 & 0 & 1 & 0 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{44} & 0 & 1 & 1 & 0 & 1 & 0 \\
1_{45} & 1 & 0 & 0 & 1 & 0 & 1 \\
1_{46} & 0 & 1 & 1 & 0 & 0 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{601} & 0 & 1 & 0 & 1 & 0 & 1
\end{pmatrix}
\end{array}
\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_7 \end{pmatrix}
$$

$$
+ \begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{44} & 0 & 0 \\
0_{45} & 1 & 0 \\
0_{46} & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{601} & 0 & 0
\end{pmatrix}_{601\times 3}
\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \\ \vdots \\ \boldsymbol{\varepsilon_{601}} \end{pmatrix}
$$

Solving the above equation gives the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ vector solutions as

$$\boldsymbol{\beta} = \begin{pmatrix} 3.3869 \\ -1.0397 \\ 0.0000 \\ 0.5391 \\ 0.0000 \\ 1.8942 \\ 0.0000 \end{pmatrix} \text{ and } \boldsymbol{\gamma} = \begin{pmatrix} 2.0215 \\ 1.0087 \\ 1.0128 \end{pmatrix}$$ for the MMPL model**,** as provided in

Table 21 and 23 respectively. For the RMPL model, the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ vector solutions respectively are

$$\boldsymbol{\beta} = \begin{pmatrix} 3.3787 \\ -1.0332 \\ 0.0000 \\ 0.5426 \\ 0.0000 \\ 1.8974 \\ 0.0000 \end{pmatrix} \text{ and } \boldsymbol{\gamma} = \begin{pmatrix} 2.0553 \\ 1.0263 \\ 1.0784 \end{pmatrix}$$

Table 21: Socio-demographic Characteristics of Heads of Household as Determinants in Dependents' Likelihood of Working Harder to Substitute for Lost Income (Model 2)

Predictor Variables	Estimation Approach							
	Maximum Pseudo-Likelihood				Residual Pseudo-Likelihood			
	B	S E	Test Statistic	P	B	S E	Test Statistic	P
Intercept	3.387	0.933	3.63	0.068	3.379	1.116	3.03	0.094
Sex								
Female	-1.04	0.433	-2.4	0.017	-1.033	0.433	-2.39	0.017
Male	0	.	.	.	0	.	.	.
Marital Status								
Single	0.539	0.403	1.34	0.181	0.543	0.403	1.35	0.179
Married	0	.	.	.	0	.	.	.
Ethnicity								
Akan	1.894	0.686	2.76	0.006	1.897	0.688	2.76	0.006
Non-Akan	0	.	.	.	0	.	.	.

the p-values of which are also provided in Tables 21 and 23 respectively, which also shows their respective *p*-values.

Now assuming $\boldsymbol{\gamma} \sim N(\mathbf{0}, \boldsymbol{G})$ and $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$, where $\boldsymbol{G}$ and $\boldsymbol{R}$ are the G–side (i.e. Gamma-side) variance-covariance matrix and R-side (Residual-side) variance-covariance matrix mentioned in Chapter 3 above, and $\mathbf{0}$ is a zero matrix, then in the model in Table 5.6, $\boldsymbol{G}$ and $\boldsymbol{R}$, for the MMPL model, are identified as

$$\boldsymbol{G} = \mathbf{I}_3\sigma_\gamma^{\mathbf{2}} = \begin{pmatrix} 2.0215 & 0 & 0 \\ 0 & 1.0087 & 0 \\ 0 & 0 & 1.0128 \end{pmatrix} \text{ and}$$

$$\boldsymbol{R} = \mathbf{I}_3\sigma_\varepsilon^{\mathbf{2}} = \begin{pmatrix} 2.1495 & 0 & 0 \\ 0 & 0.0080 & 0 \\ 0 & 0 & 0.7217 \end{pmatrix}$$

for the MMPL model**,** whereas

$$\boldsymbol{G} = \mathbf{I}_3\sigma_\gamma^{\mathbf{2}} = \begin{pmatrix} 2.0553 & 0 & 0 \\ 0 & 1.0263 & 0 \\ 0 & 0 & 1.0291 \end{pmatrix} \text{ and}$$

$$\boldsymbol{R} = \mathbf{I}_3\sigma_\varepsilon^{\mathbf{2}} = \begin{pmatrix} 3.2781 & 0 & 0 \\ 0 & 0.0083 & 0 \\ 0 & 0 & 0.7245 \end{pmatrix}$$

for the RMPL model

The significance of the G and R matrices for both MMPL and RMPL are all presented in Table 22 and 23 below.

Table 22: Covariance Parameter Estimates for Model 2

Covariance Parameter	Estimation Approach					
	Maximum Pseudo-Likelihood			Pseudo Pseudo-Likelihood		
	Subject	Est	S E	Subject	Est	S E
TYPE		2.1495	1.8517		3.2781	3.3967
AR(1)	Intercept	0.008	0.0413	Intercept	0.0083	0.0413
Residual		0.7212	0.0417		0.7245	0.042

Table 23: Solutions for Random Effects for Model 2

	Estimation Approach							
	Maximum Pseudo-Likelihood				Pseudo Pseudo-Likelihood			
Type of Household	Est	S E Pred	t Va lue	Pr > |t|	Est	S E Pred	t Val ue	Pr > |t|
HIV/AIDS	2.02	0.889	-2.27	0.023	2.06	1.081	-1.9	0.058
Other Illnesses/ Deaths	1.009	0.888	1.14	0.257	1.026	1.08	0.95	0.342
No Illness/ No Death	1.013	0.886	1.14	0.254	1.029	1.078	0.95	0.34

The variable 'Worked harder to substitute for lost household income' is the response variable which is binary in nature. The procedure specifies maximum marginal pseudo-likelihood (METHOD=MMPL) and residual marginal pseudo-likelihood (METHOD=RMPL) as the estimation techniques with a subject-specific expansion respectively. Their link functions are both logit. The "Class Level Information" table, for both models, lists the levels of the variables specified in the CLASS statement and the ordering of the levels. Seven variables are listed over there, all being explanatory variables, respectively, six of them having fixed effects and the seventh having random effect. Two of the fixed variables had three levels while the remaining four had two levels respectively. The fixed effects variable had three levels. The number of observations read and used in the analysis was 601, respectively, from the "Number of Observations" table. The size of related matrices is listed in the “Dimensions” table. The X-matrix contains 7 columns, one of which is an intercept and the remaining 6 represent the levels of the fixed effects variables all together, while the Z-matrix contains 3 columns. The random effect is made up of G-side and R-side covariance for both MMPL and RMPL models.

The “Optimization Information” table presents information about the methods and size of the optimization problem. All information was utilized by every subject in obtaining the parameters of the model given that the maximum number of observations utilized per subject as indicated by the model was 601. The optimization technique for both the MMPL and RMPL forms of the GLMM with binary data is the Newton-Raphson with Ridging, as before. The progress of the optimization process is also displayed in the “Iteration History” table. The GLIMMIX procedure performed 16 updates before the convergence criterion was met for each of the two models after the initial optimization, as displayed in the ‘Iteration history’ table. At convergence, the largest absolute value of the gradient was almost zero, indicating the fact that the process stopped at an extremum of the objective function for each of the models.

The fit of the two models is displayed in the “Model Fit Statistics” component of the SAS output and presented in Table 24. The -2Log Likelihood in the final MMPL model was 3830.42 while the -2 Residual Log Pseudo-Likelihood of the RMPL model was 3825.65. The ratio of the generalized chi-square statistic and its degree of freedom is approximately 1 for both models (i.e. 0.72 for both the MMPL and the RMPL) and measures the maximum variability in the marginal distribution of the underlying data.

Table 24: Model Fit Statistics for Model 2

Model Fit Statistics	Estimation Approach	
	Maximum Pseudo-Likelihood	Residual Pseudo-Likelihood
-2 log	3825.65	3830.42
Pseudo-AIC	3839.65	3844.42
Pseudo-AICC	3839.84	3844.61
Pseudo-BIC	3838.11	3870.39
Pseudo-CAIC	3845.11	3877.39
Pseudo-HQIC	3831.74	3851.62

The estimates and asymptotic estimated standard errors for all covariance parameters for both models are displayed in the "Covariance Parameter Estimates" table (Table 22).

The random effect, TYPE, which is estimated at 2.1495 with a standard error of 1.8517 for the MMPL model and 3.2781 with a standard error of 3.3967 for the RMPL model are found in Table 22. Again, in the “Covariance Parameter Estimates” table are found the estimates of the fixed effects, their standard errors, and their p-values.

Of all the explanatory variables utilized among the fixed effects in the MMPL model, “Gender of head of household” ($p<0.05$) and “Ethnicity” ($p<0.05$)

were the only explanatory variables that significantly explained dependents' (i.e. both children and adults above 60 years) likelihood of working harder to substitute for lost income for the household, at 95% confidence level each (Table 21). Among the fixed effects in the RMPL model, "Gender of head of household" ($p<0.05$) and "Ethnicity" ($p<0.05$) were the only explanatory variables that significantly explained dependents' (i.e. both children and adults above 60 years) likelihood of working harder to substitute for lost income for the household. However for both models, the intercept was not significant in determining respondents' likelihood of working harder to substitute for lost household income (Zhang et al., 2012).

With the G-side random effects of the models in Table 23, dependents' unlikelihood (-2.022 for the MMPL model and -2.055 for the RMPL model) of working harder to substitute for lost income varies significantly ($p=0.023$) across the HIV/AIDS category of households at 95% confidence level for the MMPL model and does not vary significantly ($p>0.05$) among the HIV/AIDS households for the RMPL model. The only significant component of the $\boldsymbol{\gamma}$ vector is 'HIV/AIDS households' with p-value equal to 0.023 (Table 23). Thus, the entire model is interpreted such that the prediction of the likelihood of occurrence of the response variable by the fixed effects is applicable to only 'HIV/AIDS households'.

For both forms of models (i.e. MMPL and RMPL), for a female-headed household, there is a less likelihood that the dependent members of the household (children under 15 years and adults above 60 years) would work harder to substitute for lost income for the household, compared to a male-headed household. In other words for households which are headed by females, the dependents are once less likely to reallocate their time to work harder to substitute for lost income in the household, all other variables held constant. That could also imply that the heads of household in such households have the capacity to pool extra incomes to make up for lost incomes, in the event of illness and death, such that there is no need for the dependent population in the household to make any further inputs in household upkeep.

Hence, in Akan-headed households, the dependent members were almost twice (1.8942 for MMPL model and 1.8974 for RMPL model) more likely to work harder to substitute for lost household income than non-Akan headed households (Table 21). However the likelihood of dependents' working harder to substitute for lost income varies across the HIV/AIDS category of households for both MMPL and RMPL models in Table 23. Thus some households in the HIV/AIDS category are more unlikely to work harder to substitute for lost income than others while that was not the case for the other two categories of households.

Dependents Leaving Job to Care for the Sick

For the response variable labelled “Dependent leaving job to care for the sick”, there are $p = 7$ fixed effects parameters to be estimated. These include Marital Status (1=single, 0=married), Ethnicity (1=Akan, 0=Non-Akan), Any recent HIV-related illness in household (1=Yes, 0=No) and the intercept component. There are also $q = 3$ random effects to be estimated. These are the conditions of the respondents in the household level, namely HIV/AIDS, Other Illnesses/Deaths and No Illness/No Death, all three being mutually exclusive.

Hence the design matrix $\boldsymbol{X}$ for this model is given as

$$\boldsymbol{X} = \begin{array}{c} \begin{array}{cccc} \text{Intercept} & \text{Marit. Stat.} & \text{Ethnicity} & \text{Recent Illness} \end{array} \\ \begin{pmatrix} 1_1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1_2 & 0 & 1 & 0 & 1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1_{44} & 1 & 0 & 1 & 0 & 1 & 0 \\ 1_{45} & 1 & 0 & 0 & 1 & 0 & 1 \\ 1_{46} & 1 & 0 & 0 & 1 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1_{601} & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}_{601\times7} \end{array}$$

The design matrix $\boldsymbol{Z}$ and response vector $\boldsymbol{y}$ are similarly respectively given as

$$\boldsymbol{Z} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1_{44} & 0 & 0 \\ 0_{45} & 1 & 0 \\ 0_{46} & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1_{601} & 0 & 0 \end{pmatrix}_{601\times3} \quad \text{and } \boldsymbol{y} = \begin{pmatrix} 0_1 \\ 0_2 \\ \vdots \\ 1_{44} \\ 0_{45} \\ 1_{46} \\ \vdots \\ 0_{601} \end{pmatrix},$$

where $1 =$ Left job to care for the sick and $0 =$ Did not leave job to care for the sick, for the MMPL model. Thus

$$
\begin{pmatrix} 0_1 \\ 0_2 \\ \vdots \\ 1_{44} \\ 0_{45} \\ 1_{46} \\ \vdots \\ 0_{601} \end{pmatrix} =
\begin{array}{cccc}
\text{Intercept} & \text{Marit. Stat.} & \text{Ethnicity} & \text{Recent Illness}
\end{array}
\begin{pmatrix}
1_1 & 0 & 1 & 0 & 1 & 0 & 1 \\
1_2 & 0 & 1 & 0 & 1 & 0 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{44} & 1 & 0 & 1 & 0 & 1 & 0 \\
1_{45} & 1 & 0 & 0 & 1 & 0 & 1 \\
1_{46} & 1 & 0 & 0 & 1 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{601} & 0 & 1 & 0 & 1 & 0 & 1
\end{pmatrix}
\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_7 \end{pmatrix}
+ \begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{44} & 0 & 0 \\
0_{45} & 1 & 0 \\
0_{46} & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{601} & 0 & 0
\end{pmatrix}_{601\times 3}
\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \\ \vdots \\ \boldsymbol{\varepsilon_{601}} \end{pmatrix}
$$

Solving the above equation gives the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ vector solution for the MMPL model as

$$
\boldsymbol{\beta} = \begin{pmatrix} 2.0048 \\ 0.5932 \\ 0.0000 \\ 0.3662 \\ 0.0000 \\ -0.3603 \\ 0.0000 \end{pmatrix} \text{ and } \boldsymbol{\gamma} = \begin{pmatrix} 0.8845 \\ 1.9392 \\ 1.0548 \end{pmatrix}, \text{ and}
$$

$$
\boldsymbol{\beta} = \begin{pmatrix} 2.0106 \\ 0.5905 \\ 0.0000 \\ 0.3705 \\ 0.0000 \\ -0.3835 \\ 0.0000 \end{pmatrix} \text{ and } \boldsymbol{\gamma} = \begin{pmatrix} 0.8995 \\ 1.9638 \\ 1.0643 \end{pmatrix}
$$

for the RMPL model. Details on their respective statistical significance are presented in Table 25.

Table 25: Determinants of the Likelihood of Dependents' Leaving their Jobs to Care for the Sick (Model 3)

Predictor Variables	Estimation Approach							
	Maximum Pseudo-Likelihood				Residual Pseudo-Likelihood			
	B	S E	Test Statistic	P	B	S E	Test Statistic	p
Intercept	2.005	0.849	2.36	0.142	2.011	1.032	1.95	0.191
Marital Status								
Single	0.593	0.254	2.34	0.02	0.591	0.254	2.33	0.02
Married	0	.	.	.	0	.	.	.
Ethnicity								
Akan	0.366	0.27	1.36	0.176	0.371	0.27	1.37	0.171
Non-Akan	0	.	.	.	0	.	.	.
Recent illness in household								
No	-0.36	0.333	-1.08	0.279	-0.38	0.334	-1.15	0.252
Yes	0	.	.	.	0	.	.	.

Assuming $\boldsymbol{\gamma} \sim N(\mathbf{0}, \boldsymbol{G})$ and $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$, where $\boldsymbol{G}$ and $\boldsymbol{R}$ are the G–side variance-covariance matrix and R-side variance-covariance matrix mentioned in Chapter 3 above, and $\mathbf{0}$ is a zero matrix, then in the model in Table 26, the matrices $\boldsymbol{G}$ and $\boldsymbol{R}$ for the MMPL model are identified as

$$\boldsymbol{G} = \mathbf{I}_3\sigma_\gamma^{\mathbf{2}} = \begin{pmatrix} 0.8845 & 0 & 0 \\ 0 & 1.9392 & 0 \\ 0 & 0 & 1.0548 \end{pmatrix}$$

and

$$\boldsymbol{R} = \mathbf{I}_3\sigma_\varepsilon^{\mathbf{2}} = \begin{pmatrix} 1.9453 & 0 & 0 \\ 0 & 0.2006 & 0 \\ 0 & 0 & 0.832 \end{pmatrix}$$

while that for the RMPL are

$$\boldsymbol{G} = \mathbf{I}_3\sigma_\gamma^{\mathbf{2}} = \begin{pmatrix} 0.8995 & 0 & 0 \\ 0 & 1.9638 & 0 \\ 0 & 0 & 1.0643 \end{pmatrix} \text{ and}$$

$$\boldsymbol{R} = \mathbf{I}_3\sigma_\varepsilon^{\mathbf{2}} = \begin{pmatrix} 2.9738 & 0 & 0 \\ 0 & 0.2003 & 0 \\ 0 & 0 & 0.8370 \end{pmatrix}$$

respectively. The significance of $\boldsymbol{G}$ and $\boldsymbol{R}$ are presented in Tables 26 and 27 below.

Table 26: Covariance Parameter Estimates for Model 3

Covariance Parameter	Estimation Approach					
	Maximum Pseudo-Likelihood			Pseudo Pseudo-Likelihood		
	Subject	Est	S E	Subject	Est	S E
TYPE		1.9453	1.661		2.9738	3.063
AR(1)	Intercept	0.2006	0.0405	Intercept	0.2003	0.0405
Residual		0.832	0.0502		0.837	0.0506

Table 27: Solutions for Random Effects for Model 3

Type of Household	Estimation Approach							
	Maximum Pseudo-Likelihood				Pseudo Pseudo-Likelihood			
	Est	S E Pred	t Val ue	Pr > |t|	Est	S E Pred	t Valu e	Pr > |t|
HIV/AIDS	0.884	0.833	1.06	0.289	0.9	1.018	0.88	0.377
Other Illnesses/ Deaths	1.939	0.84	-2.31	0.021	1.964	1.024	-1.92	0.055
No Illness/No Death	1.054	0.833	1.27	0.2	1.064	1.018	1.04	0.296

The dependent variable specified in these models is ‘Leave job to care for the sick’, with responses being ‘Yes’ (=1) and ‘No’ (=0) and is identified by PROC GLIMMIX as such, and is binary in nature. The PROC GLIMMIX estimation techniques with a subject-specific expansion, respectively, are specified by maximum marginal pseudo-likelihood (METHOD=MMPL) and residual marginal pseudo-likelihood (METHOD=RMPL). They both have ‘logit’ as their link function. The "Class Level Information" table, for both models, lists the levels of the variables specified in the CLASS statement and the ordering of the levels. Seven variables are listed, all being explanatory variables, with six of them having fixed effects and the seventh having random effect. Two of the fixed variables had three levels while the remaining four had two levels respectively. The fixed effects variable had three levels. The number of observations read and used in the analysis was 601, respectively, from the "Number of Observations" table for both MMPL and RMPL models. The dimensions of related matrices are listed in the “Dimensions” table where the X-matrix contains 7 columns, one of which is an intercept and the remaining 6 represent the levels of the fixed effects variables all together, while the Z-matrix contains 3 columns. The random effect is made up of G-side and R-side covariance parameters of dimensions 1 and 2 respectively, for both MMPL and RMPL models.

The “Optimization Information” table also presents information about the methods and size of the optimization problem, as usual. The Newton-Raphson optimization technique, with Ridging, was utilized for both MMPL and RMPL forms of the GLMM with binary data. The “Iteration History” table displays the progress of the optimization process.

The fit of the two models is displayed in Table 28. The -2Log Likelihood in the final MMPL model was 3044.54 while the -2 Residual Log Pseudo-Likelihood of the RMPL model was 3042.24.65. The ratio of the generalized chi-square statistic and its degree of freedom is approximately 1 for both models (i.e. 0.83 for the MMPL and 0.84 the RMPL) and measures the maximum variability in the marginal distribution of the underlying data.

Table 28: Model Fit Statistics for Model 3

Model Fit Statistics	Estimation Approach	
	Maximum Pseudo-Likelihood	Residual Pseudo-Likelihood
-2 log	3042.2	3044.54
Pseudo-AIC	3056.2	3058.54
Pseudo-AICC	3056.4	3058.73
Pseudo-BIC	3052.23	3087.01
Pseudo-CAIC	3059.23	3094.03
Pseudo-HQIC	3045.86	3068.21

The estimates and asymptotic estimated standard errors for all covariance parameters for both models are displayed in the "Covariance Parameter Estimates" table (Table 26).

The random effect, TYPE, which is estimated at 1.9453 with a standard error of 1.6610 for the MMPL model and 2.9738 with a standard error of 3.0630 for the RMPL model are found in Table 22. Again, in the “Covariance Parameter Estimates” table are found the estimates of the fixed effects, their standard errors, and their p-values.

Of all the explanatory variables utilized among the fixed effects in the MMPL model, only ‘Marital Status’ was significant ($p<0.05$) in predicting the likelihood of dependents leaving work to care for the sick. Similarly, among all the explanatory variables utilized among the fixed effects in the RMPL model, only ‘Marital Status’ was significant ($p<0.05$) in predicting the likelihood of dependents leaving their jobs to care for the sick The intercept of both MMPL and RMPL models were not significant in determining respondents’ likelihood of leaving their jobs to care for the sick.

The G-side random effects of the models are presented in Table 27. Here the dependents are likely to leave their jobs to care for the sick (-1.9392 for the MMPL model and –1.9638 for the RMPL model) and this varies significantly ($p=0.021$) across the ‘Other Illnesses/deaths’ category of households at 95% confidence level for the MMPL model and does not vary significantly ($p>0.05$) among the HIV/AIDS households for the RMPL model. Thus the only significant component of the $\boldsymbol{\gamma}$ vector is ‘Other Illnesses/deaths households’ with p-value equal to 0.021. This implies that with respect to the MMPL model, dependents of single-headed households with non-HIV/AIDS illnesses are more likely to leave their jobs to care for the sick than their married-headed counterparts in the event of illness in the household. With the RMPL model, there is the likelihood of the dependent members of the households to leave their jobs to care for the sick than their married-headed counterparts in the event of illness in the household. However this is not specific to any category of household.

This therefore implies that whereas the unlikelihood of dependents leaving their jobs to care for the sick in the event of illness/death varies significantly among the “Other Illnesses/Deaths” category of households, it does not vary among the “HIV/AIDS” category of households. This applies to only the MMPL model (Tables 25 and 27). Thus for some households in the “Other Illnesses/Deaths” category, dependents of single-headed households are more unlikely to leave their jobs to care for the sick than their married-headed counterparts while that is not the case for the “HIV/AIDS” and “No Illness/Death” categories of households.

Economic Determinants of Households’ Likelihood of Reducing Expenditure

For this dependent variable, there are $p = 9$ parameters to be estimated. These parameters which are the coefficients of the four relevant independent variables are Household Size (1= ‘≤5’, 0= ‘>5’), Total value of assets (1= ‘≤100’, 0= ‘>100’), Children’s education expenditure (1= ‘<100’, 0= ‘≥ 100’), Adults’ health expenditure (1= ‘≤100’, 0= ‘>100’) and intercept component.

The condition of the respondents at household level remains one of HIV/AIDS, Other Illnesses/Deaths and No Illness/No Death. Thus, $q = 3$ and all these are mutually exclusive.

Hence the design matrix $\boldsymbol{X}$ is given as

$$
\boldsymbol{X} = \begin{array}{c}
\begin{array}{ccccc} \text{Intercept} & \text{HHSize} & \text{TotValAssts} & \text{ChnEdnX'ture} & \text{AdltsHlthX'ture} \end{array} \\
\left(\begin{array}{ccccccccc}
1_1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\
1_2 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{44} & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\
1_{45} & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\
1_{46} & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{601} & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0
\end{array}\right)_{601\times 9}
\end{array}
$$

The design matrix $\boldsymbol{Z}$ and response vector $\boldsymbol{y}$ are similarly respectively given as

$$
\boldsymbol{Z} = \begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{44} & 0 & 0 \\
0_{45} & 1 & 0 \\
0_{46} & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{601} & 0 & 0
\end{pmatrix}_{601\times 3} \quad \text{and } \boldsymbol{y} = \begin{pmatrix}
0_1 \\ 0_2 \\ \vdots \\ 1_{44} \\ 0_{45} \\ 1_{46} \\ \vdots \\ 0_{601}
\end{pmatrix},
$$

where $1 =$ Household Likely to reduce expenditure and $0 =$ Household not likely to reduce expenditure, for the MMPL model. It should be noted that the vector $\boldsymbol{y}$ is not a zero vector. Thus

$$
\begin{pmatrix} 0_1 \\ 0_2 \\ \vdots \\ 1_{44} \\ 0_{45} \\ 1_{46} \\ \vdots \\ 0_{601} \end{pmatrix}
$$

$$
= \begin{array}{c}
\begin{array}{ccccc} \text{Intercept} & \text{HHSize} & \text{TotValAssts} & \text{ChnEdnX'ture} & \text{AdltsHlthX'ture} \end{array} \\
\begin{pmatrix}
1_1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\
1_2 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{44} & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\
1_{45} & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\
1_{46} & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1_{601} & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0
\end{pmatrix}
\end{array}
\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_9 \end{pmatrix}
$$

$$
+ \begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{44} & 0 & 0 \\
0_{45} & 1 & 0 \\
0_{46} & 0 & 1 \\
\vdots & \vdots & \vdots \\
1_{601} & 0 & 0
\end{pmatrix}_{601\times 3}
\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}
+ \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{601} \end{pmatrix}
$$

Solving the above equation gives the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ vector solutions as

$$
\boldsymbol{\beta} = \begin{pmatrix} 0.1976 \\ 0.2683 \\ 0.0000 \\ -0.8612 \\ 0.0000 \\ -0.4251 \\ 0.0000 \\ 1.2022 \\ 0.0000 \end{pmatrix} \text{ and } \boldsymbol{\gamma} = \begin{pmatrix} -1.6821 \\ 0.1249 \\ 1.5572 \end{pmatrix} \text{ for the MMPL model and .}
$$

$$\boldsymbol{\beta} = \begin{pmatrix} 0.1952 \\ 0.2616 \\ 0.0000 \\ -0.8530 \\ 0.0000 \\ -0.4148 \\ 0.0000 \\ 1.2111 \\ 0.0000 \end{pmatrix} \text{ and } \boldsymbol{\gamma} = \begin{pmatrix} -1.4226 \\ 0.1289 \\ 1.5937 \end{pmatrix} \text{ for the RMPL model.}$$

Detailed characteristics of the vector $\boldsymbol{\beta}$ are provided in Table 29 which also shows their respective *p*-values.

Table 29: Economic Determinants of Households' Likelihood of Reducing Expenditure as a Result of Illness/Death (Model 4)

Predictor Variables	Estimation Approach							
	Maximum Pseudo-Likelihood				Residual Pseudo-Likelihood			
	B	S E	Test Statistic	P	B	S E	Test Statistic	p
Intercept	0.198	1.169	0.17	0.881	0.195	1.324	0.15	0.896
Household size								
<= 5	0.268	0.379	0.71	0.481	0.262	0.39	0.67	0.504
> 5	0	.	.	.	0	.	.	.
Total value of assets owned (GH¢)								
<= 100	-0.86	0.388	-2.22	0.029	-0.85	0.398	-2.14	0.035
> 100	0	.	.	.	0	.	.	.
Expenditure on children's education (in GH¢)								
< 100	-0.43	0.552	-0.77	0.443	-0.42	0.566	-0.73	0.465
>=100	0	.	.	.	0	.	.	.
Health expenditure on adults (GH¢)								
< 100	1.202	0.531	2.27	0.026	1.211	0.545	2.22	0.029
>=100	0	.	.	.	0	.	.	.

Now assuming $\boldsymbol{\gamma} \sim N(\mathbf{0}, \boldsymbol{G})$ and $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$, where $\boldsymbol{G}$ and $\boldsymbol{R}$ are the G–side (i.e. Gamma-side) variance-covariance matrix and R-side (Residual-side) variance-covariance matrix mentioned in Chapter 3 above, and $\mathbf{0}$ is a zero matrix, then in the model in Table 22, $\boldsymbol{G}$ and $\boldsymbol{R}$, for the MMPL model, are identified as

$$\boldsymbol{G} = \mathbf{I}_3\sigma_\gamma^{\mathbf{2}} = \begin{pmatrix} 1.6821 & 0 & 0 \\ 0 & 0.1249 & 0 \\ 0 & 0 & 1.5572 \end{pmatrix}$$

$$\boldsymbol{R} = \mathbf{I}_3\sigma_\varepsilon^{\mathbf{2}} = \begin{pmatrix} 1.8597 & 0 & 0 \\ 0 & 0.0445 & 0 \\ 0 & 0 & 0.6627 \end{pmatrix}$$

For the MMPL model

$$\boldsymbol{G} = \mathbf{I}_3\sigma_\gamma^{\mathbf{2}} = \begin{pmatrix} 1.7226 & 0 & 0 \\ 0 & 0.1289 & 0 \\ 0 & 0 & 1.5937 \end{pmatrix}$$

$$\boldsymbol{R} = \mathbf{I}_3\sigma_\varepsilon^{\mathbf{2}} = \begin{pmatrix} 2.8963 & 0 & 0 \\ 0 & 0.0467 & 0 \\ 0 & 0 & 0.6985 \end{pmatrix}$$

for the RMPL model. The significance of the G and R is presented in Tables 30 and 31 below.

Table 30: Covariance Parameter Estimates for Model 4

	Estimation Approach					
Covariance Parameter	Maximum Pseudo-Likelihood			Pseudo Pseudo-Likelihood		
	Subject	Est	S E	Subject	Est	S E
TYPE		1.8597	1.64		2.896	3.058
AR(1)	Intercept	0.0445	0.12	Intercept	0.047	0.116
Residual		0.6627	0.10		0.699	0.103

Table 31: Solutions for Random Effects for Model 4

Type of Household	Estimation Approach							
	Maximum Pseudo-Likelihood				Pseudo Pseudo-Likelihood			
	Est	S E Pred	t Value	Pr > |t|	Est	S E Pred	t Value	Pr > |t|
HIV/AIDS	1.682	0.841	-2	0.05	1.723	1.03	-1.67	0.098
Other Illnesses/ Deaths	0.125	0.835	0.15	0.88	0.129	1.024	0.13	0.9
No Illness/ No Death	1.557	0.839	1.86	0.07	1.594	1.027	1.55	0.124

The variable 'Reducing expenditure' is entered in the PROC GLIMMIX procedure as a response variable with binary distribution. The procedure specifies maximum marginal pseudo-likelihood (METHOD=MMPL) and residual marginal pseudo-likelihood (METHOD=RMPL) as the estimation techniques with a subject-specific expansion respectively in the "Model Information" table in the SAS output. They both possess logit link functions. The "Class Level Information" table, for both models, lists the levels of the variables specified in the CLASS statement, in the procedure, and the ordering of the levels. Here, five explanatory variables are listed, four of which have fixed effect and have two levels each while the eight has a random effect, and has three levels.

Information about the methods and size of the optimization problem are contained in the "Optimization Information" table in the SAS output and illustrated in the appendix. The optimization technique for both the MMPL and RMPL forms of the GLMM with binary data is the Newton-Raphson with Ridging. The progress of the optimization process is also displayed in the

"Iteration History" table, where the GLIMMIX procedure performed only 12 iterations before the convergence criterion was met for the first model (i.e. MMPL). However for the second model (i.e. RMPL), there were 13 iterations before convergence could be reached, as displayed in the 'Iteration history' output of the SAS PROC GLIMMIX procedure. At convergence, the largest absolute value of the gradient was almost zero, indicating the fact that the process stopped at an extremum of the objective function for each of the models.

The fit of the two models is illustrated in the "Model Fit Statistics" component of the SAS output and presented in Table 22. The -2Log Pseudo-Likelihood in the final MMPL model was 416.46 while the -2 Residual Log Pseudo-Likelihood of the RMPL model was 413.43. The ratio of the generalized chi-square statistic and its degree of freedom was 0.66 (close to 1) for the MMPL model and 0.70 (closer to1) for the RMPL model. These represent the measure of their respective levels of variability in the marginal distributions of the underlying data. This implies that the MMPL model presents data, for this dependent variable, that are close to each other with respect to the response variable. It further implies that for the MMPL model, the likelihood of reducing or not reducing expenditure is closer among households than it is the case for the RMPL model. In other words, for the MMPL model, the households have more common effects as a result of changes in the independent variables than it is the case for the RMPL model.

Table 32: Model Fit Statistics for Model 4

Model Fit Statistics	Estimation Approach	
	Maximum Pseudo-Likelihood	Residual Pseudo-Likelihood
-2 log	413.43	416.46
Pseudo-AIC	431.43	434.46
Pseudo-AICC	433.55	436.44
Pseudo-BIC	426.35	454.42
Pseudo-CAIC	435.35	463.42
Pseudo-HQIC	418.15	440.72

The estimates and asymptotic estimated standard errors for all covariance parameters for both models are displayed in the "Covariance Parameter Estimates" table in Table 32. The random effect, TYPE is estimated at 1.8597 with a standard error of 1.6445 for the MMPL model while that of the RMPL is estimated at 2.8963 with a standard error of 3.0575 also in Table 32. Furthermore, in the "Covariance Parameter Estimates" table in the SAS output of the PROC GLIMMIX procedure, illustrated in the appendix, are found the estimates of the fixed effects, their standard errors and their p-values (Table 32). Of all the explanatory variables utilized among the fixed effects in the MMPL model, "Health expenditure on adults (GH¢)" ($p<0.05$) and "Total value of assets owned (in GH¢)" ($p<0.05$) were the only significant explanatory variables that explained households' likelihood of reducing expenditure as a result of illness and/or death, whereas among the fixed effects in the RMPL model, "Health expenditure on adults (GH¢)" ($p<0.05$) and "Total value of assets owned (in GH¢)" ($p<0.05$) were the only significant explanatory variables that explained households' likelihood of reducing expenditure as a result of illness and/or death. The remaining explanatory variables (namely “household size”, “Self-described economic status in community” and Expenditure on child’s education”) were not significant in determining households’ likelihood of reducing expenditure in the face of illness and/or death in both models.

With respect to the G-side random effects, in Table 33, labelled “Solutions for Random Effects”, the likelihood of reducing expenditure varies significantly among the HIV/AIDS households for the MMPL model (-1.682 for the MMPL model, $p<0.05$). It did not however vary significantly for the RMPL mode. In vector terms, $\gamma_1 =$ HIV/AIDS households’ is the only significant component of the $\boldsymbol{\gamma}$ vector ($p<0.05$), the entire model is interpreted such that the prediction of the likelihood of occurrence of the response variables by the fixed effects is applicable to only ‘HIV/AIDS households’ for the MMPL model.

Thus on the fixed effects, for the MMPL model, households which have total value of assets to the tune of less than GH¢100.00 per month, are less likely to reduce expenditure compared to households whose total value of assets is to the tune of GH¢100.00 or more, all other factors held constant. This also applies to the RMPL model. For the RMPL model as well, households which have total value of assets to the tune of less than GH¢100.00 per month, are less likely to reduce expenditure compared to households whose total value of assets is to the tune of GH¢100.00 or more, all other factors held constant. This simply implies that the lower the total value of household assets, the less likely that household is to reduce expenditure and vice versa, for both models.

On the other hand, the higher the health expenditure incurred on a household, the more likely that household is to reduce all other expenditure in general, and vice versa. This applies to both models.

Households' likelihood of reducing expenditure (-1.682 for the MMPL model and -1.723 for the RMPL) varies significantly ($p<0.05$) among the "HIV/AIDS" households for the MMPL model but does not vary significantly ($p>0.05$) among the "HIV/AIDS" households for the RMPL model. This implies that, for the MMPL model, the HIV/AIDS category of households employ very similar coping strategies in the face of illness or death.

Discussion of Further Analysis

It is noted that PROC GLIMMIX does not build models directly based on the original data, unlike other SAS model building procedures (Wolfinger and O'Connell, 1993; Breslow and Clayton 1993). PROC GLIMMIX rather builds models on what is called "pseudo-data". It begins the whole iteration process for each model by constructing a new data (pseudo-data) based on the "real" or original data (Wolfinger and O'Connell 1993) and this is done by "linearizing" the original data around the expected values emanating from the earlier or previous iteration (Broström, 2003; Schabenberger and Gregoire ,1996; Pinheiro and Bates, 2000; SAS ,2010; Wolfinger and O'Connell , 1993). It then maximizes the pseudo-likelihood on that pseudo-data. This is referred to as the outer iteration (Wolfinger and O'Connell, 1993; Schabenberger and Gregoire, 1996).

The fixed and random effects parameters are not solved all at the same time in the algorithm for GLMM in Proc Glimmix, as it is for other models (and software). Instead, it begins the iteration with a pseudo-likelihood, which is itself a function of variance components, fixed effects, and the dispersion coefficient and is able to tease out the fixed effects and dispersion coefficient. In the end it uses these to generate an objective function which, in itself, is a function of the variance components alone (Kieman et al., 2012). It continues by going further into another level of iteration in order to optimize the earlier modified objective function over the earlier derived variance components only. Together with the generated estimates of the variance components from the earlier iteration it, estimates fixed effects and predicts its accompanying random effects. The process then goes back to the outer iteration for another round of iteration by producing another set of pseudo-data set and the process continues until convergence is finally achieved, i.e. where the difference in parameter estimates between successive linear mixed model fits falls within a specified tolerance level (Broström, 2003; Schabenberger and Gregoire ,1996; Pinheiro and Bates, 2000).

The afore-mentioned explains why in GLMMs, the fixed and random effects are partitioned into X (matrix) and Z (matrix) terms respectively. The partitioning therefore gives GLMM an edge over the other forms of Linear Models (LMs) in that beyond the fixed effects, the random effects (of GLMMs) provide further information as to how these effects vary among members in a group (or levels of the random effects). These are very necessary if solutions should be provided to problems to which these models are applied. The random effects help in tailoring solutions or interventions to respective groupings or clusters within the population and not the entire population being fitted with a wholesale solution (or intervention) which in this regard works for some segments of the population and not others. It is to this end that Mzolo et al.'s (2009) recommendation in their study that the correlation observed at an enumeration area (EA) level indicates that interventions should consider the area effect (i.e. grouping or cluster effect) rather than only the individuals (i.e. the whole population altogether). They continued by making reference to Grosskurth (1995) who also recommended in their Mwanza Trial that studies that intervene at the community level should be encouraged to fight diseases such as HIV and TB. They used the community (EA) level as their random factor. This edge is what was absent in the study of Pitayanon et al. (1998). However Grosskurth's (1995) work was not about identifying the determinants of impact of illnesses and/or deaths on the dependent members of the household.

In the case of both pseudo-AIC and pseudo-BIC values, a smaller value indicates a better model fit (Arrandale, 2006). This therefore means that in the bid to, for instance, compare the MMPL model to the RMPL model, the stronger model must, as a matter of necessity, have the values of all the fit statistics including the pseudo-AIC and pseudo-BIC being smaller. Where one is smaller and the other is larger, the models become incomparable. Applying this to our data in this study, Table 5.4 shows that whereas the Pseudo-AIC (3103.7) is less than the Pseudo-BIC (3096.7) for the MMPL model, the counterparts are larger for the RMPL (Pseudo-AIC=3106.6, Pseudo-BIC=3151.9) model. This therefore makes the MMPL model the more preferable, in our scenario (Arrandale et al., 2006). The outcomes of the two models appear very similar. For instance, for the MMPL and RMPL models in Table 5.1, only one variable (i.e. “Recent illness in household”) among the fixed effects variables was responsible for reallocation of time (the outcome variable) among the dependents in the household, and this varies across households which experienced illnesses and/or deaths other than HIV/AIDS-related ones only. This implies that in the event of an illness with an adult household member, the dependents of non-HIV/AIDS households are forced to reallocate their time. However, the impact is felt in some households more than others.

Unlike the case of the dependents in the non-HIV/AIDS households where some were more affected than the others, dependents in “HIV/AIDS” were equally affected with respect to reallocation of time in the event of recent illness. However, those dependents in “No Illness/No Deaths” households were also affected, not necessarily because there was a recent illness or death in their households, but probably because of some other extraneous factor(s) other than illness and death. This presents a more detailed picture of how the effect of recent illness affects the reallocation of the time of dependents in the population. This picture is what was absent in the study of Pityanon et al. (1997) and Mzolo et al. (2011). They only presented their findings in terms of the entire population in a wholesale manner, though they explained the circumstances restraining them from delving deeper.

This finding, of recent illness in the household being responsible for reallocation of the time of the dependent members in the household, is very similar to the scenario cited in Balyamujura et al. (2000) where a young boy, who should be a dependent, but whose mother was sick in the Zambia said he spent much time looking for money by doing menial jobs in order to raise some money to cater for his/her sick mother. This Zambian case is a classic example of reallocation of dependent household members' time as a result of illness and/or death. Rugalema (1999a) also has it that households severely affected by illnesses exhaustively resort to child labour as a coping mechanism. Rugalema (1999b) further emphasized that the illness affects time allocation.

The implications of this finding, in the case of this study, is that interventions for this affected population will be tailored differently for the households experiencing other illnesses and deaths rather than as against those of the other two ("HIV/AIDS" and "No Recent Illness/Deaths"), who will need a common intervention since the impact is the same among and across them. Drimie (2002) brought to the fore the fact that it is important to recognize the variability in the impact of HIV/AIDS on rural households. He said the poorer households, especially those with small land holdings are much less able to cope with the effects of HIV/AIDS than wealthier households who can hire casual labour and are better able to absorb shocks. This view perfectly falls in line with the random components of this model which has it that impact varies across levels of the random factors.

On the model fit statistics in Table 5.8, the Pseudo-AIC (3839.65) is greater than the Pseudo-BIC (3838.11) for the MMPL model, while the Pseudo-AIC (3844.42) is less than the Pseudo-BIC (3870.39) in the RMPL model. Hence Hence as in Model 1, the MMPL is the preferred model. The models in Table 5.5 also provide evidence to the effect that dependents in female-headed households were less likely to work harder to substitute for lost household income than in households headed by males, and this applies to both MMPL and RMPL models. Moreover this finding varies across only HIV/AIDS-related households. In households experiencing other illnesses or deaths apart from HIV as well as those even experiencing no recent illnesses or deaths, the experiences of dependents of female headed households did not vary.

One revealing finding about this study is the fact that dependents in Akan-headed households were almost twice more likely to work harder to substitute for lost household income than those of non-Akan-headed ones. This finding supports several earlier studies regarding the impacts of HIV/AIDS on the dependent households, (Balyamujura et al., 2000). They indicate that within the household setting, the female heads were more affected, and hence more vulnerable, than male heads as far as coping mechanisms were concerned. This has been largely published in numerous articles (Donahue et al., 2000; Walker, 2002). Walker has it that reasons why women are more vulnerable than men to HIV/AIDS, for instance, include female physiology, women's lack of power to negotiate sexual relationships with male partners, especially in marriage, and the gendered nature of poverty, with poor women particularly vulnerable (Walker, 2002), and that comes to bear in this study. Indeed in most of the literature it has been largely held that much of the burden of fending for the households has arisen as a result of the death of the male heads (Waterhouse and Vifjhuizen, 2001).

Interestingly, Akan-headed households were about twice more likely to work harder to substitute for lost household income than the non-Akan headed ones. Very little, if any, is known about tribe factor in determining what factors are responsible for dependent members of households working harder to substitute for lost income of those households. The UNAIDS earmarked three major areas where strategies to cope with the HIV/AIDS pandemic can be categorized and these have been referred to by Balyamujura et al. (2000). These include strategies aimed at improving food security, raising supplementary income to maintain the household expenditure patterns and those aimed at alleviating the loss of labour. These are clearly in consonance with some of the findings of this study, in that dependents of female-headed households were found to be less likely to work extra hard to support the household as opposed to their counterparts in male-headed households. This is explained by Loewenson and Whiteside (1997), who intimate that women frequently carry a double burden of generating income outside the home and for care giving as well as maintaining family land, in the case of HIV/AIDS. From here it could be deduced that these women would work extra hard to take care of the work that children would have done to support within the household setting.

The implications of this finding, in the case of this study, is that interventions for this affected population would have to be tailored differently for the HIV/AIDS households as against those of the other two ("Other Illnesses/Deaths" and "No Recent Illness/Deaths") households, who will need a common intervention since the impact is the same among and across them.

The MMPL model in Table 5.12 is more preferred than its RMPL counterpart. This stems from the fact that while for the MMPL model, the Pseudo-AIC was 3056.2, the Pseudo-BIC was 3052.23; whereas in the RMPL model, the Pseudo AIC was 3058.54 while the Pseudo-BIC was 3087.01 (Anderson Burnham and Thompson, 2000). Furthermore, in Table 5.9, dependents of female-headed households were more likely to leave jobs to care for the sick, in both MMPL and RMPL models, than those of their male-headed counterparts, and this is about twice less likely to vary among households experiencing illnesses and deaths other than HIV/AIDS. It does not however vary among households experiencing HIV/AIDS and those not experiencing any recent illnesses or deaths. This study also brings another dimension of the impact of illnesses and/or deaths on the dependent members of the household, which is hardly seen in other literature, to the fore. This stems from the facts that not just are female heads of household very likely to leave their jobs to care for the sick, but their dependents as well. This impact is very variable among households experiencing illness and/or deaths other than those of HIV/AIDS. This could be attributed to their vulnerability to most of the harsh social factors that hit very hard at them in the event of illness and/or death of the male heads of households (Walker, 2002).

The Pseudo-AIC and Pseudo-BIC for the MMPL model in Table 5.16 are 431.43 and 426.35 respectively for the MMPL model, while they are 434.46 and 454.42 respectively for the RMPL model. This again makes the MMPL model the more preferred than its RMPL counterpart. Variables used in constructing the models in Table 5.9 through Table 5.16 include "household size", "Total value of assets owned", "Expenditure on Children's education" and "Health expenditure on adults", only "Total value of assets owned" and "Health expenditure on adults" were the significant predictors of households' likelihood of reducing expenditure as a result of illness and/or death. Furthermore, households owning assets up to GH¢ 100.00 were less likely to reduce expenditure as a result of illness and/or death, compared to those owning more than GH¢ 100.00, for both MMPL and RMPL models. Also, households incurring health expenditure on adults to the tune of less than GH¢ 100.00 were less than one-and-a-half times more likely to reduce expenditure as a result of illness and/or death, compared to those incurring GH¢ 100.00 or more, and this applies to both MMPL and RMPL models. Furthermore, these findings are almost twice less likely to vary among the HIV/AIDS households for both MMPL and RMPL models (per the G- and R-side random effects of the model). They were also more likely to vary among the "No Illness/Death" category of households for only the MMPL models. Interestingly, there was no significant variability among the "No Illness/Death" category of households for only the RMPL model. For the "Other Illnesses/Deaths" category of households as well, there was no significant variability of the findings for both MMPL and RMPL models.

On the reduction of household expenditure as a result of illness and/or death, Rugalema (1999a) in the Tanzania study found that short and long-term costs resulted in reduction in the household expenditure through curtailing the number and quality of meals that a household could afford which resulted in poor nutrition with obvious implications for health. These could have been the means by which expenditure was reduced to be spent on health in our study as well. Following from this study as well, it turns out that households owning more valuable assets were rather more likely to reduce expenditure than those owning less valuable assets. This could explain why Cohen (1993), HSRC (2001a) and Rugalema (1999a) showed that HIV/AIDS first affects the welfare of households through illness and death of family members, which in turn leads to the diversion of resources from savings and investments into care.

On a vertical comparison of the four models (from Table 5.1 through Table 5.16) used in this study based on the "Smaller is better" principle, and the principle of parsimony (Daniels et al., 1999, Anderson, 1973, Fan, Huang and Li, 2007), whereas all the Pseudo-AIC values were less than their Pseudo-BIC counterparts for the RMPL method of all the four models, they (Pseudo-AIC values) were less than their Pseudo-BIC counterparts in three out of four models in the RMPL method. Furthermore, given that the RMPL models consistently turned out the larger pseudo-AICs throughout, the MMPL method of modeling happens to be better than the RMPL.

A number of interesting findings were made regarding HIV/AIDS and gender in this study. It was discovered for example that in female-headed households, there is less likelihood of the dependent population (children under 15 years and adults above 60 years) to work harder to substitute for lost income resulting from the death of a member. Dependents of female-headed households were more likely to leave their jobs to care for the sick, in both MMPL and RMPL models, than those of male-headed households.

The gender findings of this study have important implications. For example, it could mean, in one breath, that females are very capable of providing for their families, and consequently do not need the dependent population of their households to make further inputs into household upkeep. It could also mean, in another breath, that extra pressure is put on female heads as they have to shoulder the financial burden of caring for their entire families alone. It can safely be concluded then that female heads are more affected by HIV/AIDS illness/deaths at the household level than male heads, and are therefore more vulnerable than their male counterparts. It is therefore recommended that HIV/AIDS interventions be directed adequately at female populations in HIV/AIDS households. That is not to say that males do not deserve attention in HIV/AIDS interventions, but only to suggest that females present a heightened need which requires urgent attention. To ensure the success of interventions, it is recommended that HIV/AIDS interventions have a gender angle to them, to enable them proactively address the needs of females in fighting HIV/AIDS.

Analysis of the educational qualifications of respondents from HIV/AIDS households also provides interesting insights. It was discovered that the highest proportion of respondents (65.5%) in HIV/AIDS households held

Middle/JHS/SHS qualifications. The second largest group of respondents (12.2%) among "HIV/AIDS" households with respect to the highest level of education attained was those with Primary education. The type of household with the highest number of respondents with no education was the HIV/AIDS household (4.4%). The educational dynamic has huge implications for HIV/AIDS prevention and management in Ghana.

The higher a person moves up the educational ladder, the greater access he or she has to knowledge and information. In other words, higher educational attainment brings higher levels of exposure to information. Drawing from this, it is safe to argue that educated individuals are more likely to gain access to HIV/AIDS information than less-educated individuals. If this is true, then concerted efforts need to be channeled into increasing the flow of HIV/AIDS information to HIV/AIDS households. The particularly low-to-average level of educational attainment in these households presents a possibility for restricted access to HIV/AIDS information. To stem the propensity for the development of an HIV-infection cycle in these households, targeted efforts should be directed at supplying more HIV/AIDS information to members of HIV/AIDS households. This will empower surviving members of these households and prevent them from also getting infected, thereby halting a cycle of HIV/AIDS infection in the households.

It was discovered that Akans as an ethnic group dominate in "HIV/AIDS" and "Other illnesses/deaths" households at 38.1% and 38.6% respectively. This implies that to an extent, Akans are affected more by HIV/AIDS than other ethnic groups. But it must be quickly added that the Akan population, from the 2010 Population Census, makes about half of the total population of Ghana. In the Greater Accra alone, it is more than half of the population of Ghana. This could be the reason for this outcome. Hence if this study should be replicated in other regional capitals, the story is very likely going to be different regarding ethic group.

The study revealed that members of HIV/AIDS households had to significantly reallocate their time as a result of illness/death. A substantial proportion of HIV/AIDS households needed to find jobs (14.1%), work harder to substitute for lost income (12.6%), and leave school for work (11.9%). Also, when key household members (major breadwinners) die, other dependent members are forced to find other means of survival. It was discovered for example that 56.8% of HIV/AIDS households borrowed money as a result of sickness/death of a member compared to only 12.5% of “Other Illnesses” households. This goes to show that HIV/AIDS has a more intense impact on households than other diseases.

It is evident from this study that a major reason for the phenomenon of ‘child labour’ in this country is the illness/death of a key (breadwinner) member of a household. It was discovered that 11.9% of members of HIV/AIDS households leave school for work. This category of ‘school leavers’ includes children as well. Though this statistic pertains to HIV/AIDS households, it is not limited to them alone. Other households record similar situations when key members of those households become inactive as a result of illness or death.

Summary

The study's main objective is to conduct a review of the GLMMs and apply it to a surveyed primary data on the effect that the occurrence of diseases and deaths have on the household. Two types of model have been determined for a number of variables on coping strategies in terms of some suitable independent variables. The two types of model are the MMPL and the RMPL. The RMPL model indicates whether estimation is based on residual likelihood of the mean of the random effects using a Pseudo-Likelihood technique whereas the MMPL indicates whether estimation is based on maximum likelihood of the mean of the random effects using a Pseudo-Likelihood technique. The coping strategy variables examined in this chapter are Sex, Age, Marital Status, Recent illness in household, Ethnicity, Household size, Total value of assets owned by household (GH¢) and Health expenditure on adults (GH¢). Each of these was utilized in the model as relevant independent variables.

The fixed effects utilized in the Generalized Linear Mixed Models happened to be the coping strategies while the random effect was "type of household". The first model was "determining the fixed and random effects of the predictors of dependents' likelihood of reallocating their time". In this model "Recent illness in household" was the only fixed effect that predicted dependents' likelihood of reallocating their time. This finding was applicable to only "HIV/AIDS" category of households, with respect to the only random effect in the model.

The second model was “predicting determinants of dependents working harder to substitute for lost household income”. The fixed effects for this model were Sex, Marital Status and Ethnicity while the random effect was “Type of household”. In the end, “Gender of head of household” and ethnicity were the only fixed effects that could predict the dependents working harder to substitute for lost household income and this applies to only HIV/AIDS category of households.

The third model sought to determine the fixed and random effects that were responsible for “Dependents Leaving Job to Take Care of the Sick”, with fixed effects being Marital Status, Ethnicity, and “Any recent HIV-related illness in household” and the random effect was Type of Household. Marital Status was the only fixed effect that was able to significantly determine Dependents likelihood of Leaving Job to Take Care of the Sick and this was applicable to the “Other Illnesses/Deaths” categories of household.

In the fourth model, “Economic Determinants of Households’ Likelihood of Reducing Expenditure”, the fixed effects used were Household Size, Children’s Education Expenditure, Total Value of Assets Owned (in GH¢) and Adults’ Health Expenditure and the random effect was Type of Household. However, “Health expenditure on adults (GH¢)” and “Total value of assets owned (in GH¢)” were the only significant fixed effects and these were applicable to the “HIV/AIDS” category of households in the random effect “Type of Household”.

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATIONS

Introduction

The summary of the study is presented in this section, followed by the discussion and conclusion in the second and third sections respectively, while the recommendations are presented in the fourth section.

Summary

The study's main objective is to conduct a review of the GLMMs and apply it to primary data on the effect that the occurrence of diseases and deaths have on the household. The right hand side of the Generalized Linear Mixed Models (GLMMs) is made up of the fixed and random components. The fixed components conform to the usual linear models while the random components, also known as random effects, identify the disaggregation of the variability of the fixed effects of the model into subgroups in the target population. The random effect is further identified by the $\boldsymbol{G}$ and $\boldsymbol{R}$ covariance matrices. The Maximum Mean Pseudo-Likelihood (MMPL) and the Residual Mean Pseudo-Likelihood (RMPL) were the two among several approaches at estimating the coefficients of the random components of the GLMM which was applied in this study.

Two types of model were determined for a number of variables on coping mechanisms adopted by households in the event of these illnesses and deaths. The two types of model studied are the MMPL and the RMPL. The RMPL model indicates whether estimation is based on residual likelihood of the mean of the random effects using a Pseudo-Likelihood technique whereas the MMPL indicates whether estimation is based on maximum likelihood of the mean of the random effects also using a Pseudo-Likelihood technique.

The background of this study was presented in Chapter One. In Chapter Two, Key variables associated with illnesses and eventual deaths within households were listed. These formed the bases for the data collected in this study. Three categories of household were studied. These are households with an adult member experiencing a recent HIV or AIDS episode, also referred to as "HIV/AIDS household", households where an adult member has a recent episode of diseases other than HIV/AIDS, called "Other illnesses Households" and the third being households which had no recent illness or death experience, also referred to as "No illnesses/deaths household". These three categories of household studied formed the three levels of the random effect of each of the four models developed in this study.

The fixed component of the GLMM was made up of variables such as occupation, religion, value of household assets, number of children cared for, health, education and upkeep expenditure incurred on children and that incurred on adults in terms of health and upkeep, among others.

In the preliminary analysis, HIV/AIDS households were found to be more predominantly headed by females than the other two categories. The heads of the HIV/AIDS category of households were a shade older than their counterparts from the other two categories. The HIV/AIDS households incurred more deaths than the “Other Illnesses/Deaths” households. Very few of the HIV/AIDS category of households owned assets such as house, farm land, building land, a car, livestock, etc., compared to their counterparts in the other two categories of household. The HIV/AIDS category of households had the least total value of assets owned compared to the other two categories. Most of HIV/AIDS households also had more dependents than the other two categories of household.

The HIV/AIDS households incurred more expenditure on their dependent children’s health than their education and upkeep, contrary to what pertains in the other two categories of household. They also incurred far less total expenditure on their dependent children than the other two categories of household before the onset of the disease but not after. Then HIV/AIDS category of households also incurred a lower cost of medical treatment, monthly income loss and funeral expenses than their “Other Illness/Deaths” category of households. However, they incurred a higher travel cost than their counterparts from the “Other Illness/Deaths” category of households.

There was further analysis, where the GLMM was applied to real life data, collected for this purpose, in order to determine which fixed factors could predict the impact of illness and death on households. The variable “Type of illness”, with three levels, was used as the random effect of the model. Four models were fitted with dependent variables being “dependents re-allocating their time as a result of illness or death of an adult household member”, “dependents working harder to substitute for lost income as a result of the illness or death of an adult household member", "dependents leaving their jobs to care for the sick" and "Households' likelihood of reducing expenditure as a result of illness and death".

The first of the four models was "determining the fixed and random effects of the predictors of dependents' likelihood of reallocating their time". In this model "Recent illness in household" was the only fixed effect that predicted dependents' likelihood of reallocating their time. This finding was applicable to only "HIV/AIDS" category of households, with respect to the only random effect in the model.

The second model was "predicting determinants of dependents working harder to substitute for lost household income". Sex, Marital Status and Ethnicity while the random effect was "Type of household" were the fixed effects, with their respective levels. "Gender of head of household" and "ethnicity" were the only fixed effects that could predict "dependents working harder to substitute for lost household income" and this applies to only HIV/AIDS category of households.

The third model was aimed at determining the fixed and random effects that were responsible for "Dependents Leaving Job to care for the Sick", with the fixed effects being Marital Status, Ethnicity, and "Any recent HIV-related illness in household" and the random effect was Type of Household. For this model, "Marital Status" was the only fixed effect that was able to significantly determine Dependents likelihood of Leaving Job to care for the Sick and this was applicable to the "Other Illnesses/Deaths" categories of household.

"Economic Determinants of Households' Likelihood of Reducing Expenditure" was the fourth model. The fixed effects used in this model were "Household Size", "Children's Education Expenditure", "Total Value of Assets Owned (in GH¢)" and "Adults' Health Expenditure" and the random effect was "Type of Household". "Health expenditure on adults (GH¢)" and "Total value of assets owned (in GH¢)" were the only significant fixed effects and these were applicable to the "HIV/AIDS" category of households in the random effect "Type of Household".

Conclusion

"A recent HIV/AIDS-related illness in the household" is the only factor (or random effect) responsible for dependents' likelihood of re-allocating their time as a result of illness/death. In the MMPL model, it applies to only HIV/AIDS households however with the RMPL model the random effect is not significant in predicting re-allocation of dependents' time using PROC GLIMMIX.

"Sex" and "Ethnicity" are the only factors that are responsible for the likelihood of dependents' leaving job to care for the sick as a result of illness/death, and this applies to only HIV/AIDS households per the MMPL and does not apply to any of the three categories of household per the RMPL model using PROC GLIMMIX.

Marital status is the only factor that is responsible for the likelihood of dependent members of the household leaving job to care for the sick as a result of illness/death and this applies to households with other illnesses of deaths due to other illnesses per the MMPL model. It does not however apply to any of the three categories of households per the RMPL model using PROC GLIMMIX.

"Total value of assets owned" is the only factor (or fixed effect) that is responsible for the likelihood of households reducing expenditure as a result of illness/death and this applies to only households with other illnesses or deaths due to other illnesses per the MMPL model. It does not apply to any of the three categories of households per the RMPL model using PROC GLIMMIX.

It also turned out that for the MMPL model, the fixed factors which were significant in predicting the dependent variables were also applicable to the random factors whereas for the RMPL model, the fixed factors which were significant in predicting the dependent variables were not applicable to the random factors.

It was discovered from this study that recent illnesses and deaths within the households of the target population impact the dependents by causing reallocation of the dependents' time. However, whereas this impact varies across households with other Illnesses or deaths due to other illnesses, it does not vary among households where HIV/AIDS-related illnesses and deaths occurred in the past one year. Furthermore, even within households where no recent illnesses or deaths took place, the study provided evidence that they also encountered other extraneous factors that caused the dependents to reallocate their time apart from those considered in the model.

The study provides evidence to the effect that dependents of female-headed households were less likely to work harder to substitute for lost household income than those of male-headed households. In another model, it turns out that dependents of female-headed households were more likely to leave their jobs to care for the sick than those of their male-headed counterparts, and this is less likely to vary among households that belong to the "other illnesses/Deaths". Also, it turns out that Akan-headed households were about twice more likely to work harder to substitute for lost household income.

Adults' Health Expenditure and Total Value of Assets Owned by the household were the only significant economic determinants of households' likelihood of reducing expenditure. These were applicable to only HIV/AIDS category of households according to the MMPL model and not by the RMPL model.

Recommendations

Indeed, households that experience HIV/AIDS-related illnesses or deaths should be given more attention as far as social support is concerned than those with other illnesses and deaths emanating from them.

The study reveals that female heads of household are less likely to work harder to substitute for lost income. This could be due to the fact that they are always working extra hard to fend for the household. Hence female-headed households need more social support in the event of illness or death, particularly the HIV/AIDS-headed households. Also, Akan-headed households need more social support that non-Akan-headed households in the event of illness or death of a household member.

In light of the revealing and salient findings made in this study, the following recommendations are made for consideration by policy makers.

Interventions targeted at HIV/AIDS need to be gender-specific. It was discovered that in Akan-headed households, the dependent members of the household were almost twice (1.8942 for MMPL model and 1.8974 for RMPL model) more likely to work harder to substitute for lost household income compared to non-Akan headed households. These particular phenomena are peculiar, and reasons for this need to be investigated in further research. It would be helpful to know if the Akan situation is due to cultural and attitudinal factors, or some other accentuating factor. This can greatly aid in addressing HIV/AIDS and other diseases at the ethnic level. It is therefore recommended that sociological and anthropological research be undertaken into the effects of tribal factors in HIV/AIDS impact on households.

It is also recommended that credit/financial programs be instituted to provide some financial support to HIV/AIDS affected households. These schemes will extend funds to HIV/AIDS households to augment their meager incomes. This will go a long way to ameliorate the effects of the disease on their finances. It would also serve as a buffer against the onset of vicious cycles of poverty which tend to marginalize and further entrench the already-unbearable situations of these households.

It is recommended that public education and awareness be intensified to address social, cultural and individual deficiencies in dealing with HIV/AIDS and other illnesses and diseases.

To ameliorate the effects of HIV/AIDS and other illnesses on affected households, it is recommended that social support programs be instituted to prop up affected households. A key area that these interventions should be directed at is educational attainment. The programs should be able to ensure, for example, that children from affected households stay in school and continue to acquire an education in order to have a secure future.

It is recommended that future studies consider several interaction effects among the factors studied in this study (Drimie, 2002). This will enable deeper insights into the interplay of factors considered in this study, and how they can further be managed to ensure enhanced livelihoods for households.

Further studies should be carried out into the ethnic group factor in order to gain further and deeper insight into disease impact upon the household.

REFERENCES

Agatha, C. O., Walingo, M. A., & Othuon, L. (2010). Expenditure Patterns On Food And Non-Food Items In Hiv/Aids Affected And Non-Affected Households In Kisumu District, Kenya. *African Jurnal of Food, Agriculture and Development (AJFAND)*.Vol. 10, No. 4.

Agresti, A. (2002), *Categorical Data Analysis*, Second Edition, New York: John Wiley & Sons.

Agyepong, I. A., & Kangeya-Kayonda, J. (2004). Providing Practical Estimates Of HIV/AIDS Burden For Health Planners In Resource-Poor Countries, *Am J Trop Med Hyg.*

Ainsworth, M., Beegle, K., & Koda, G. (2005). The impact of adult mortality and parental deaths on primary schooling in Northwestern Tanzania. *Journal of Development Studies*, 41(3), pp.412 – 439.

Akeroyd, A. V. (1997). 'Sociocultural aspects of AIDS in Africa: occupational and gender issues.' in Bond G C, Kreniske J, Susser I & Vincent J (eds) *AIDS in Africa and the Caribbean.* Westview, Oxford

Anand, S., & Hanson, K. (1997). Disability-Adjusted life years: a critical review. Journal of Health Economics 16 (1997) 685-702

Anderson, T. W. (1973). Asymptotically efficient estimation of covariance matrices with linear structure. Ann. Statist. 1 135–141. MR0331612

Ansell, N., & Young, L. (2004). "Enabling households to support successful migration of AIDS orphans in Southern Africa." AIDS Care **16**(1): 3.

Arrandale, V. H. (2006). An Evaluation of Two Existing Methods for Analyzing Longitudinal Respiratory Symptom Data [M.Sc. Thesis]. Vancouver: University of British Columbia.

Jankovic S, Vlajinac H, Bjegovic V, Marinkovic J, Sipetic-Grujicic S, Markovic-Denic L, Kocev N, Santric-Milicevic M, Terzic-Supic Z, Maksimovic N, Laaser U. (2007). The burden of disease and injury in Serbia. The burden of disease and injury in Serbia. *Eur J Public Health* 17(1):80-5

Attanayake N., Fox-Rushby J., & Mills A. (2000). Household costs of 'HIV/AIDS' morbidity: a study in Matale district, Sri Lanka. *Trop Med Int Health 9:* 595–606

Balyamujura, H., Jooste, A., van Schalkwyk, H., Geldenhuys, F., Crew, M., Carstens, J., Bopape, L. & Modiselle, D. S. (2000). 'The Impact of HIV/AIDS on Agriculture', unpublished research report for the National department of Agriculture, University of the Free State (Chair in International Agricultural marketing & Development), University of Pretoria (Centre for AIDS) and the Agricultural Research Council, South Africa, October.

Barnett, T. & Blaikie, T. P. (1992). *AIDS in Africa: Its Present and Future Impact*. London: Belhaven Press.

Barnett, T. & Haslwimmer, M. (1995). *The Effects of HIV/AIDS on Farming Systems in Eastern Africa*. Rome: Food and Agriculture Organization of the United Nations.

Barnett, T. & Blaikie, P. (1992a) *Simple Methods for Monitoring the Socio-Economic Impact of AIDS: Lessons from Research in Uganda and Kenya, Development Studies Discussion Paper* no. 223, School of Development Studies, University of East Anglia, September.

Barnett, T., Whiteside, A. & Desmond, C. (2001). The Social and Economic Impact of HIV/AIDS in Poor Countries: A Review of Studies and Lessons. Progress in Development Studies 1 (2): 151-170.

Bates, I., Fenton, C., Gruber, J., Lalloo, D., Lara, A.M., Squire, S.B., Theobald, S., Thomson, R. & Tolhurst, R., (2004). Vulnerability to malaria, tuberculosis, and HIV/AIDS infection and disease. Part 1: determinants operating at individual and household level. THE LANCET Infectious Diseases Vol. 4, May 2004.

Bicego G., Rutstein, S., Johnson, K., (2003). "Dimensions of the emerging orphan crisis in sub-Saharan Africa." *Social Science and Medicine* 56:1235-1247.

Bloom, D. E. & Canning, D. (2003).The Health and Poverty of Nations: from Theory to Practice. Journal of Human Development, 4 (1),

Bollinger, L., Stover, J. & Antwi, P. (1999). The Economic Impact of AIDS in Ghana The Futures Group International in collaboration with Research

Triangle Institute (RTI), The Centre for Development and Population Activities (CEDPA).

Booysen, F. l. R. (2002). "Financial Responses of Households in the Free State Province to HIV/AIDS-Related Morbidity and Mortality." *South African Journal of Economics* 70(7): 1193-1215.

Borkow, G., Weisman, Leng, Z., Stein, Q., Kalinkovich, M., Wolday, A., Bentwich, D. & Helminths, Z. (2001). Human Immunodeficiency Virus and Tuberculosis. Scandanavian Journal of Infectious Diseases, 2001, 33: 568-571.

Breslow, N. E. & Chatterjee, N. (1999). Design and analysis of two-phase studies with binary outcome applied to Wilms tumour prognosis. *Applied Statistics* 48(4), 457-468

Breslow, N. E. & Clayton, D. G., (1993). "Approximate Inference in Generalized Linear Mixed Models," *Journal of the American Statistical Association*, 88, 9–25.

Broström, G. (2003). Technical Reports. Umeå, Sweden: Department of Statistics, Umeå University. Generalized linear models with random intercepts. SE-90187

Brouwer, C. N., Lok, C. L., Wolffers, I. & Sebagalls, S. (2000). “Psychosocial and economic aspects of HIV/AIDS and counselling of caretakers of HIV-infected children in Uganda.” *AIDS Care*. 12: 5, 535-540.

Burton, P. R., (2003). Correcting for nonrandom ascertainment in generalized linear mixed models (GLMMs), fitted using Gibbs sampling. *Genet Epidemiol* 24:24–35, 2003. 2003 Wiley-Liss, Inc.

Burton, P. R., Tiller K. J., Gurrin, L. C., Cookson, W. O., Musk, A. W. & Palmer, L. J. (1999). Genetic variance components analysis for binary phenotypes using generalized linear mixed models (GLMMs) and Gibbs sampling. Genetic epidemiology 17:2 pg 118-40

Carballo, M. & Siem, H. (1996) Migration Policy and AIDS. In Crossing Borders: Migration Ethnicity and AIDS *Taylor & Francis, London, pp. 31–49.*

Case, A., Paxton, C. & Ableidinger, J. (2004) Orphans in Africa: Parental Death, Poverty, and School Enrolment. *Demography,* 41, 483-508

Cohen, B. & Trussell, J. (1996). Preventing and Mitigating AIDS in Sub-Saharan Africa:

Cohen, D. & Reid, E. (1998). "The vulnerability of women: Is this a useful construct for policy and programming?" *UNDP HIV and Development Programme, Issues Paper* No 28, http://www.undp.org/hiv/issue28E.htm

Cohen, D., (1993). "The HIV Epidemic in Uganda: a Programme Approach,' *UNDP HIV and Development Programme*, *Study Paper* No. 1, http://www.undp.org/hiv/study1E.htm

Cohen, D. (1998). Poverty and HIV/AIDS in Sub-Saharan Africa. HIV and Development Programme. Issue Paper no. 27. 1998.

Collins, J. & Rau, W. (2000). ‘AIDS in the context of development’, United Nations Research Institute for Social Development (UNRISD) Programme on Social Policy and Development, Paper No. 4, December.

Crampin, A. C., Floyd, S., Glynn, J. R., Madise, N., Nyondo, A., Khondowe, M. M., ... Fine, P.E. (2003). The long-term impact of HIV and orphanhood on the mortality and physical well-being of children in rural Malawi. *AIDS*. 17(3):389-97.

Cross, C., (2001). Sinking deeper down: HIV/AIDS as an economic shock to rural households. *Society in Transition*, Vol. 32(1). 133-147

Daniels, M. J. & Kass, R. E. (1999). Nonconjugate Bayesian estimation of covariance matrices and its use in hierarchical models. J. Amer. Statist. Assoc. 94 1254–1263. MR1731487

Danziger, R. (1994). “The social impact of HIV/AIDS in developing countries.” *Social Science and Medicine* 39(7): 905-17.

De Savigny, D. & Binka, F. (2004). “Monitoring Future Impact on Malaria Burden in Sub- Saharan Africa”. *American Journal of Tropical medicine and Hygiene. 71 (2suppl), 224-31.*

De Waal, A. & Whiteside, A. (2003), 'New Variant Famine’: AIDS and Food Crisis in Southern Africa, *The Lancet*, 362.

Deininger, K., Garcia, M. & Subbarao, K. (2003). "AIDS-induced orphanhood as a systemic shock: magnitude, impact, and program interventions in Africa." *World Development* 31(7):1201-1220.

Diggle, P. J., Liang, K. Y. & Zeger, S. L. (1994). Analysis of Longitudinal Data.Oxford: Oxford University Press.

Dixon-Fyle, K. & Mulanga, C. (2004). Responding to HIV/AIDS in the world of work in Africa: the role of social protection. Working Paper 5. ILO Programme on HIV/AIDS and the world of work. Geneva, June 2004.

Dobson, A. (1990). *An Introduction to Generalized Linear Models.* Chapman and Hall, London.

Donahue, J., Kabbucho, K. & Osinde, S. (2000) 'HIV/AIDS – Responding to a silent economic crisis among microfinance clients', unpublished report, MicroSave Africa.

Dorrington, R., Bourne, R., Bradshaw, D., Laubscher, R. & Timaeus, I.M. (2001). The impact of HIV/AIDS on adult mortality in South Africa. South African Medical Research Council, Cape Town.

Drimie, S. (2003). HIV/AIDS and Land: Case Studies from Kenya, Lesotho and South Africa. Development Southern Africa, December 2003. 20(5).

Ellis, F. (2000). *Rural Livelihoods and Diversity in Developing Countries*, Oxford: Oxford University Press

Ettling, M. B. & Shepard, D. S., (1991). Economic cost of HIV/AIDS in Rwanda. *Trop Med Parasitol 42:* 214–218.

Eze, J. I. (2000). Modelling HIV/AIDS Epidemic in Nigeria. A Doctoral Dissertation submitted to the Department of Statistics, University of Glasgow.

Fan, J., Huang, T. & Li, R. (2007). Analysis of longitudinal data with semiparametric estimation of covariance function. J. Amer. Statist. Assoc. 102 632–641. MR2370857

FAO HIV/AIDS Programme (2002) ‘HIV/AIDS, food security and rural livelihoods’, available at: www.fao.org.

Foster, G. & Williamson, J. (2000). "A Review of Current Literature of the Impact of HIV/AIDS on Children in Sub-Saharan Africa." *AIDS* 14 (Supplement 3): S275S284.

Foster, G. (1998). “Today’s Children—Challenges to Child Health Promotion in Countries with Severe AIDS Epidemics.” *AIDS Care* 10 (Suppl 1): S17-S23.

Gallup, J. L., Sachs & J. D. (2001). The economic burden of HIV/AIDS. *Am J Trop Med Hug 64 (Suppl 1–2):* 85–96

Ghana News Agency (GNA, 2009). 160,000 Ghanaian children made orphans through HIV/AIDS. News Article extracted from http://news.myjoyonline.com/health/200903/28156.asp , Grant, M. R., & Palmiere, A. D. (2003). When Tea is a Luxury: The Economic Impact of HIV/AIDS in Bulawayo, Zimbabwe. African Studies 2003. 62(2).

Gregson, S., Garnett, G. P. & Anderson, R. M. (1994). “Assessing the Potential Impact of the HIV-1 Epidemic on Orphanhood and Demographic Structure of Populations in sub-Saharan Africa.” *Population Studies* 48(3): 435-458

Grosse, S. D., Donald J. L., Campbell, V. A. & Chamie, M. (2009). Disability and Disability-Adjusted Life Years: Not the Same. Public Health Reports. Vol. 127.

Grosskurth, H., Todd, J., Mwijarubi, E., Klokke, A., Senkoro, K., Mayaub, P., Changaluch, J., Nicoll, A. & Ka-Gina, G. (1995). Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomized controlled trial. *Lancet*, 346 (8974), 530-536.

Gross, J. (1997). More AIDS is seen in people over 50. *New York Times.*

Grundfest-Schoepf, B. (1991). ”Ethical, methodological and political issues of AIDS research in Central Africa,‘ *Social Science and Medicine*, 33(7), pp 749-763

Hagen, J. (2002). "A generation of orphans”: The General Assembly Debates.

Hopkins D. R., Richards F. O. & Katabarwa, M. (2005). Editorial: Whither Onchocerciasis Control in Africa? M. *Am. J. Trop. Med. Hyg.*, 72(1), pp. 1–2

Hosmer, D. W. & Lemeshow, S. (2000). *Applied Logistic Regression*. John Wiley & Sons, Inc., Toronto.

Hunger**,** M**.,** Döring, A. & Holle, R. (2012). Longitudinal beta regression models for analyzing health-related quality of life scores over time. *BMC Medical Research Methodology*, 12:144

Hunter, Susan S. (1991). "The Impact of AIDS on Children in Sub-Saharan African Urban Centers." *African Urban Quarterly* 6(1 &2):108-128.

International Fund for Agricultural Development (2001). 'Strategy paper on HIV/AIDS for East and Southern Africa', working document. Human Sciences Research Council (2001a) 'HIV/AIDS in Southern Africa: A review paper for the W. K. Kelloggs Foundation', unpublished report by the Social Aspects of HIV/AIDs and Health Programme for W. K. Kelloggs, Foundation, October.

Jooma, B. M. (2005). Southern Africa Assessment: Food Security and HIV/AIDS. African Security Review 2005.14(1).

Jankovic, S.. (2005). Summary measures of population health and their relevance for health policy. In: Galan A, Scintee G, editors. Public Health Strategies. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company;. p.190-207. Available from URL: http://www.snz.hr/phsee/publications.htm. Accessed: July 31, 2009

Kamali, A. Seeley, J. A. Nunn, A. J. Kengeya-Kayondo, J. F. Ruberantwari, A. & Mulder, D. W. (2010). The orphan problem: experience of a sub-Saharan Africa rural population in the AIDS epidemic. AIDS Care 8(5):509-515.

Kareem, S. A., Raviraja, S., Awadh, N. A., Kamaruzaman, A., & Kajindran, A. (2010). Classification And Regression Tree In Prediction Of Survival Of AIDS Patients. *Malaysian Journal of Computer Science*, 23(3).

Khan, M. J. (1966). Estimate of economic loss due to HIV/AIDS in West Pakistan. *Pakistan J Health 16:* 187–193.

Kiszewski A. E. & Teklehaimanot A. (2004). *A Review Of The Clinical And Epidemiologic Burdens Of Epidemic HIV/AIDS,* Am J Trop Med Hyg

Kleinman, K., Lazarus, R. & Platt, R. (2014). A generalized linear mixed models approach for detecting incident clusters of disease in small areas, with an application to biological terrorism. *Am J Epidemiol. 2004 Feb 1; 159(3):217-24.*

Knodel, J. & VanLandingham, M. (2003). “Return migration in the context of parental assistance in the AIDS epidemic: The Thai experience.” *Social Science and Medicine* 57(2):327-342.

Knodel, J., Saengtienchai, C. Im-Em, W. & VanLandingham, M. (2001). "The impact of AIDS on parents and families in Thailand—A key informant approach." *Research On Aging* 23 (6): 633-670.

Kominski, G. F., Simon, P. A., Ho, A., Luck, J., Lim, Y. W. & Fielding, J. E. (2002). Assessing the burden of disease and injury in Los Angeles County using disability-adjusted life years. *Public Health Rep*;117:185-91.

Koper, N. & Manseau, M. (2010). A guide to developing resource selection functions from telemetry data using generalized estimating equations and generalized linear mixed models. *Rangifer*, Special Issue No. 20: 195–203

Kwaramba, P. (1997). "The socio-economic impact of HIV/AIDS on communal agricultural systems in Zimbabwe." Working Paper # 19. Harare: Friedrich Eibert Stiftung/Zimbabwe Farmers Union.

Lau, J. T. F. & Wong, W. S. (2001). AIDS-related Discrimination in the Workplace: The Results of Two Evaluative Surveys Carried Out During a Three-Year Period in Hong Kong. *AIDS CARE* 2001. 13(14): 433-440.

Llanos, A., Hertrampf, E., Cortes, F., Pardo, A., Grosse, S. D., Uauy, R. (2007). Cost-effectiveness of a folic acid fortification program in Chile. *Health Policy*;83:295-303

Lipson, D. (2001). GATS and trade in health insurance services: background note for WHO Commission on Macroeconomics and Health, CMH working paper series paper no WG4: 7. June, 2001. Geneva: World Health

Organization,2001.

http://www.whoindia.org/LinkFiles/Commision_on_Macroeconomic_and_Health_04_07.pdf (accessed Dec 01, 2011)

Loevinsohn, M. & Gillespie, S. (2003). HIV/AIDS, Food Seciruty, and Rural Livelihoods: Understanding and Responding. *IFPRI, FCND Discussion Paper* No. 157.

Loewenson, R. & Whiteside, A. (1997). 'Social and economic issues of/HIV/AIDS in Southern Africa', report prepared for SAFAIDS, Harare, Occasional Paper Series, No. 2.

Loewenson, R. & Chikumbirike, T. (2005). Community Based Monitoring and Research on Food Security and Social Welfare. Paper for the SARPN meeting, South Africa, May 2005.

Lugalla, J. P. L., Emmelin, M. A. C., Mutembei, A. K., Comoro, C.J., Killewo, J. Z. J., Kwesigabo, G., Sandstrom, A. I. M. & Dahlgren, L. G. (1999). The Social and Cultural Contexts of HIV/AIDS Transmission in the Kagera Region, Tanzania. Koninklijke Brill NV, Leiden. 377-402.

Magadi, M. (1992). "The Impact of AIDS on Some Urban Households in Kenya with Particular Reference to Female Headed Households" *African Urban Quarterly* 7(3/4):285-8.

Mathers, C. D., Vos, E. T., Stevenson, C. E. & Begg, S. J. (2001). The burden of disease and injury in Australia. Bull World Health Organ;79:1076-84.

Matshalaga, N. R.. (2002). "Social Dynamics of Orphan Care in the Era of the HIV/AIDS Pandemic: An Insight of Grandmothers' Experiences in Zimbabwe." Dissertation Abstracts: *The Humanities and Social Sciences* 63 (3):1143-A.

McCullagh, P. & Nelder, J. A. (1989). *Generalized Linear Models*, Second Edition, London: Chapman & Hall.

McKenzie, F. E., Samba, E. M. (2004). The role of mathematical modeling in evidence- based malaria control. *American Journal of Tropical Medicine and Hygiene* 71s2: 94-96.

Melse, J. M., Essink-Bot, M. L., Kramers, P. G. & Hoeymans, N. A. (2000). national burden of disease calculation: Dutch disability-adjusted life-years. Dutch Burden of Disease Group. *Am J Public Health*;90:1241-7.

Michaud, C.M., McKenna, M.T., Begg, S., Tomijima, N., Majmudar, M., Bulzacchelli, M. T., Ebrahim, S., Ezzati, M., Salomon, J.A., Kreiser, J.G., Hogan, M., Murray, C.J. (2006). The burden of disease and injury in the United States, 1996. *Popul Health Metr*;4:11.

Monasch. R., & Boerma, J. T. (2004). Orphans and child care patterns in sub-Saharan Africa: an analysis of 40 national surveys. AIDS; 18 (Suppl. 2):S55–S65.

Munyako, D. (1994). "The Impact of Structural Adjustment Programs on the Health of Kenyan Women" Mimeo.

Murray, C. J. L., & Lopez, A. D. (2004). Assessing health needs: the Global Burden of Disease Study. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford Textbook of Public Health - Fourth edition. New York: Oxford University Press Inc. p.243-54.

Myers, R. H. (1990). *Classical and Modern Regression with Applications*. PWS-KENT, Boston, 2nd edition.

Mzolo, T., (2011). Investigating factors associated with HIV among 15 to 24 year old females. tUL Diepenbeek.

Mzolo, T., Mwanbi, H. & Zuma K. (2009). Bayesian versus Frequentist Approaches in Estimating Risk Determinants of Infectious Diseases. http://www.statssa.gov.za/isi2009/ScientificProgramme/IPMS/1241.pdf (Sourced on August 25, 2014)

NACP (2011). National HIV Prevalence and AIDS Estimates Report.

Nattrass, N. (2004). The Moral Economy of AIDS in South Africa. United Kingdom: Cambridge University Press.

Neter, J., Wasserman, W., & Kutner, M. H., (1985). *Applied linear statistical models* (2nd ed.). Homewood, IL: Irwin.

Ngalula, J., Urassa, M., Mwaluko, G. R., Isingo, J. & Ties Boerma J. (2002). "Health service use and household expenditure during terminal illness due to AIDS in rural Tanzania." *Tropical Medicine & International Health* 7(10): 873-877.

Niehof, A. (2004). A micro-ecological approach to home care for AIDS patients. *Medische Antropologie* 16 (2).

Niehof, A. 2005. The Significance of Diversification for Rural Livelihood Systems. Food Policy 29: 321-338.

Ntozi, J. P. M., Anarfi, J. K., Caldwell, J.C., & Jain, S. (eds.) (1997). ”Vulnerability to HIV infection and effects of AIDS in Africa and Asia/India,‘ *Health Transition Review*, Supplement to Volume 7, Health Transition Centre, The Australian National University, Canberra Australia, 1997.

Ntozi, J. P. M., & Nakanaabi, I. M. (1997), ”The AIDS epidemic and infant and child mortality in six districts of Uganda,‘ *Health Transition Review*, Supplement to Volume 7, pp 189205

Ntozi, J. (1997). “Effect of AIDS on children: the problem of orphans in Uganda.” *Health Transit Rev.* 7(suppl.):23-40.

Nyambedha E. O., Wandibba, S. & Aagaard-Hansen, J. (2001). “Policy implications of the inadequate support systems for orphans in western Kenya.” *Health Policy* 58:83

Nyambedha E. O., Wandibba, S. & Aagaard-Hansen, J., (2003). “Changing patterns of orphan care due to the HIV epidemic in western Kenya.” *Social Science and* Medicine 57:301-311.

Oppong, Y. (2001). Public Health and Profitability: Perceptions and Responses of the Business Community in Ghana to HIV/AIDS and Malaria In Collaboration with Belfer Center for Science and International Affairs John F. Kennedy School of Government, Harvard University. African EconomicPolicy Discussion Paper No. 78, August 2001. Funded by USAID, Washington, DC.

Pinheiro, J.C. & Bates, D. M., (2000). Mixed-effects models in S and S-PLUS. New York: *Springer*.

Piot, P., Bartos, M., Ghuys, P.D., Walker, N. & Schwartlander, B. (2001). The Global Impact of HIV/AIDS. Nature 2001. Vol. 410.

Pitayanon, S., Kongsin, S. & Janjaroen, W. (1997). The economic impact of HIV/AIDS mortality on households in Thailand. In The economics of HIV and AIDS: The case of South and South East Asia, ed. D. Bloom, and P. Goodwin. Delhi: *Oxford University Press*.

Population Studies Series (2004). The Impact of Aids. No. 229, United National Department of Economic and Social Affairs.

Powers D. A. & Xie, Y. (1999). Statistical Methods for Categorical Data Analysis. Academic Press, Inc.

Preble, E.A. (1990). “Impact of HIV/AIDS on African children.” *Social Science and Medicine* 6:671-680.

Pridmore, P. & Yates, C. (2005). "Combating AIDS in South Africa and Mozambique: the role of open, distance, and flexible learning (ODFL)." Comparative Education Review 49(4): 490

Pridmore, P. (2008). Access to conventional Schooling for \Children and Young People Affected by HIV and AIDS in Sub-Saharan Africa: A Cross national review of recent research evidence. *IOE, London.* Research and Data Priorities for the Social and Behavioural Sciences. National Academy Press, Washington, DC.

Rosegrant, M. W., Paisner, M. S., Meijer, S. & Witcover, J. (2001). 2020 Global Food Outlook: Emerging Trends, and Future Alternatives. IFPRI, Washington.

Rugalema, G. (1999a). ‘Coping or struggling? A journey into the impact of HIV/AIDS on rural livelihood in Southern Africa’, unpublished paper.

Rugalema, G. (1999b). ‘It is not only the loss of labour: HIV/AIDS, loss of household assets and household livelihood in Bukoba District, Tanzania’, in Mutangadura, G; Jackson, H & Mukurazita, D, (eds) *AIDS and African Smallholder Agriculture*, Southern African AIDS Information Dissemination Service (SAFAIDS), Harare.

Russell, S. (2004). The economic burden of illness for households in developing countries: catastrophic or manageable? A review of studies focusing on HIV/AIDS, tuberculosis, and human immuno-deficiency virus/acquired immunodeficiency syndrome. *Am J Trop Med Hyg 71 (Suppl 2)*, 147–155

Ryder, R.W., Kamenga, M., Nkusu, M., Batter, V. & Heyward, W.L. (1994). AIDS orphans in Kinshasa, Zaire: incidence and socioeconomic consequences. *AIDS*. 8:673-679.

Samatebele, H. (2005). Civil Society Organisation (CSO) Experiences of Monitoring of Food Security Issues in Zambia. A Paper presented at the Southern African Regional Poverty Network (SARPN) Regional Workshop, 24th May 2005, Johannesburg, South Africa.

SAS Institute, (2010). The GLIMMIX Procedure: SAS Publishing.

Sauerborn, R., Shepard, D. S., Ettling, M. B., Brinkmann, U., Nougtara, A., Diesfeld, H. J. (1991). Estimating the direct and indirect economic costs of HIV/AIDS in a rural district of Burkina Faso. *Trop Med Parasitol 42:* 219–223

Save the Children, (2004). Children in a world of AIDS. Extracted from http://www.savethechildren.org/atf/cf/%7B9def2ebe-10ae-432c-9bd0-df91d2eba74a%7D/World_of_AIDS_1004.pdf on 12/12/2010

Schabenberger, O., Gregoire, T. G. (1996). Population-averaged and subject-specific approaches for clustered categorical data. *Journal of Statistical Computation and Simulation*.;54:231–253.

Schackman, B.R., Neukermans, C.P., Fontain, S.N., Nolte, C., Joseph, P., Pape, J.W., Fitzgerald & D.W. (2007). Cost-effectiveness of rapid syphilis screening in prenatal HIV testing programs in Haiti. *PLoS Med*;4:e183.

Schachterle, S. E., Mathes, R. W., Paladini, M. & Weiss, D. (2013). Long-Term Asthma Trend Monitoring in New York City: A Mixed Model Approach. Online J Public Health Inform. 2013; 5(1): e123.

Sengendo, J., & Nambi, J. (1997). The psychological effect of orphanhood: a study of orphans in Rakai district. *Health Transit Rev. 1997;7 Suppl:105-24. Unique Identifier : AIDSLINE MED/97421468*

Shelton, J. D., Cassell, M. M. & Adetunji, J. (2005). Is Poverty or Wealth at the Root of HIV? *The Lancet*. Vol. 366.

Shepard, D. S., Ettling, M. B., Brinkmann, U., & Sauerborn, R. (1991). The economic cost of HIV/AIDS in Africa. *Trop Med Parasitol 42:* 199–203

Shisana, O. & Simbayi, L. (2002). South African National HIV Prevalence, Behavioural Risks and Mass Media. Household Survey 2002. HSRC, Cape Town.

Sibanda, E.N., Stanczuk, G. & Kasolo, F. (2003). International Archives of Allergy and Immunology, 2003. Vol. 132: 183-195.

Sida, (1998b). HIV in the World Today: A Summary of Trends and Demographic Implications.

Sinha, A., Levine, O., Knoll, M. D., Muhib, F. & Lieu, T. A. (2007). Cost-effectiveness of pneumococcal conjugate vaccination in the prevention of child mortality: an international economic analysis. Lancet;369:389-96

SPSS Inc. (1993-2007). 233 South Wacker Drive, 11th Floor Chicago, IL 60606-6412.

Stewart, J. (2003). Moving Food: The World Food Programme's Response to the Southern African Humanitarian Crisis. African Security Review, 2003. 12(1) :17-27.

Stokes, M. E., Davis, C. S., & Koch, G. G., (2000). *Categorical Data Analysis Using the SAS System*, Second Edition, Cary, NC: SAS Institute Inc.

Subbarao, K., Mattimore., A., & Plangemann, K. (2001). "Social Protection of Africa's Orphans and Other Vulnerable Children: Issues and Good Practice Program Options." Africa Region Human Development Working Paper Series. Human Development Sector, Africa Region, The World Bank.

The Abuja Declaration and the Plan of Action, April 25, (2000). Source: (http://mosquito.who.int/docs/abuja_declaration.pdf), accessed on Mar 12, 2011.

Topouzis, D. & Hemrich, G. (1994). The Socio-economic Impact of HIV and AIDS on rural families in Uganda: An Emphasis on Youth. *UNDP Study Paper* No. 2

Topouzis, D. (2003). "The Implications of HIV/AIDS for Rural Development Policy and Programming: Focus on Sub-Saharan Africa." Study Paper No. 6. Geneva: UNDP HIV and Development Program. Available at:

Toy, M., Oguz, F., Wörmann, T., Bozdayi, A. M., Schalm S. W., Borsboom, G. J., van Rosmalen Richardus, J. H. & Yurdaydin, C. (2011). Age- and region-specific hepatitis B prevalence in Turkey estimated using generalized linear mixed models: a systematic review. BMC Infectious Diseases. http://www.biomedcentral.com/1471-2334/11/337. Accessed on 04092014.

Ugwuanya, E. B. (2003). The Impact of Aids on Intergenerational Relationships in Nigeria: The Position of the Aged. *Miami University and OhioLINK*

UNAIDS (1998c). Epidemiological Fact Sheet on HIV/AIDS and Sexually Transmitted Diseases: ëUgandaí, UNAIDS/WHO, Geneva.

UNAIDS (1999) 'A review of household and community responses to the HIV/AIDS epidemic in the rural areas of Sub-Saharan Africa, UNAIDS, Geneva, Switzerland.

US Bureau of the Census, (1998). HIV/AIDS Surveillance Database, Population Division, International Programs, January

Uys, D. & Leana, R. (2003). "Aspects of the Care of People with HIV/AIDS in South Africa." *Public Health Nursing* 20(4): 271-280.

VanLandingham, M., Knodel J., Im-em, W. & Saengtienchai C. (2000). The Impacts of HIV/AIDS on Older Populations in Developing Countries: Some Observations based upon the Thai Case. Population Studies Center, University of Michigan. *Research Report* No. 99-441

Verbeke, G. (2006). Mixed models for the analysis of categorical repeated measures. *PAGE*, Bruges.

Verner, D. & Alda, E. (2004). Youth at Risk, Social Exclusion, and Intergenerational Poverty Dynamics: A New Survey Instrument with Application in Brazil. World Bank Policy Working Paper 3296, May 2004.

Wachter K. W., Knodel, J. E. & VanLandingham, M. (2002). “AIDS and the elderly of Thailand: projecting familial impacts.” *Demography* 39(1):25-41.

Walker, C. (2002). ‘Ensuring women’s land access’, moderators comments, Regional Workshop on Land Issues in Africa and the Middle East, Kampala, Uganda, April 29 – May 2, 2002.

Walraven, G., Nicoll, A., Njau, M. & Timaeus, I. (1996). "The impact of HIV-1 infection on child health in sub-Saharan Africa: the burden on the health services." *Tropical Medicine and International Health* 1(1): 3-14.

Warwick, I., Bharat, S., Castro, R., Garcia, R., Leshabari, M.T., Singhanetra-Renard, A. & Aggleton, P. (1998). Household and community responses to HIV and AIDS in developing countries. Critical Public Health, 8(4).

Waterhouse, R. & Vifjhuizen, C. (2001). ‘Strategic women and gainful men: gender, land and natural resources in different rural contexts in Mozambique’, Nucleo de Estudos de Terra, UEM; Faculty of Agronomy and Forestry Engineering, UEM and ActionAid – Mozambique, Maputo.

Whiteside, A. (2002). "Poverty and HIV/AIDS in Africa." *Third World Quarterly* 23(2): 313-332.

Wiggins, S. (2005). Southern Africa's Food and Humanitarian Crisis of 2001-04: Causes and Consequences. Discussion Paper, Agriculture Economic Society Annual Conference, Nottingham.

Wijngaarden, J. & Shaeffer, S. (2005).The impact of HIV/AIDS on children and young people, reviewing research conducted and distilling implications for the education sector in Asia: HIV/AIDS & education. Bangkok: UNESCO Bangkok, 14 p. (Discussion paper, no. 1).

Wolfinger, R. & O'Connell, M. (1993). "Generalized Linear Mixed Models: A Pseudo-likelihood Approach," *Journal of Statistical Computation and Simulation*, 4, 233–243.

World Health Organization, (2002). "Impact of AIDS on Older People in Africa: Zimbabwe Case Study." Geneva: WHO Ageing and Life Course Programme. Available at: http://www.eldis.org/static/DOC11093.htm.

Wyss, K., Hutton, G. & N'Diekhor, Y. (2004). Costs attributable to AIDS at household level in Chad. *AIDS CARE*. October 2004, 16(7), 808-816.

Zaba, B. & Gregson, S. (1998). "Measuring the impact of HIV on fertility in Africa." *AIDS* 12 (Suppl 1): S41-50.

Zabaa, B., Whiteside, A. & Boerma, T. J. (2004). Demographic and socioeconomic impact of AIDS: Taking stock of the empirical evidence. AIDS, 18 (suppl 2):S1–S7.

Zhang, X., Zhang, Y., Aleong, T., Baker, T., & Fuller-Thomson, E. (2012). Factors associated with the Household Income of Persons Living with HIV/AIDS in China. Global Journal of Health Science 4(3).

Zuccolo, L., Milena M., Maule, & Dario G. (2005). An Epidemiological Application of a Bayesian Nonparametric Smoother Based on a GLMM with an Autoregressive Error

APPENDICES

Appendix A

Variables used in Modeling

	Original Coding	New Coding
Sex		
	Male	Male
	Female	Female
Age of respondent		
	15-34	15-34
	35-49	35-49
	50+	50+
Current Marital Status		
	Single	Single
	Married	Married
Type of Household		
	HIV/AIDS	HIV/AIDS
	Other Illnesses/Deaths	Other Illnesses/Deaths
	No Illness/No Deaths	No Illness/No Deaths
Ethnicity		
	Akan	Akan
	Ewe	Non-Akan
	Ga	
	Northerner	
	Other	

Appendix A (Continued)

Variables used in Modeling

	Original Coding	New Coding
Highest level of education completed		
	None	Up to JSS/JHS
	Primary	SHS
	Middle/JHS/SHS	Post Sec & Others
	Voc./Comm.	
	Tertiary	
	Other	
Recent illness in household		
	No	No
	Yes	Yes
Household Size		
	1 to 5	≤ 5
	6 to 10	> 5
	15 or more	
Self-described economic status		
	Very well off	Well-off
	Well off	Poor
	Moderately well off	
	Poor	
	Very poor	

Appendix A (Continued)

Variables used in Modeling

Original Coding	New Coding
Health Expenditure on children's education	
None	< 100
<100	≥ 100
100 – 999	
1,000 and above	
Health Expenditure on children's health	
None	< 100
<100	≥ 100
100 and above	
Upkeep Expenditure on adults	
Not Applicable	< 100
< 100	≥ 100
100 – 999	
1,000 - 9,999	
10,000 and above	
Health Expenditure on adults	
Not Applicable	< 100
< 100	≥ 100
100 – 999	
1,000 - 9,999	

Appendix B

Survey Questionnaire

INFORMED CONSENT

Hello. My name is DAVID YAO MENSAH and I am a student at the Department of Mathematics and Statistics at the University of Cape Coast. I am writing my PhD thesis on A review of estimation techniques in generalized linear mixed models with application to disease impact modelling. I would very much appreciate your participation in this survey. The information you provide will help me establish the extent to which morbidity and mortality incidences in households affect members of those households. The survey usually takes between 15 and 25 minutes to complete. Whatever information you provide will be kept strictly confidential and will not be shown to other persons.

Participation in this survey is voluntary and you can choose not to answer any individual question or all of the questions. However, I hope that you will participate in this survey since your views are important.

At this time, do you want to ask me anything about the survey?

May I begin the interview now?

Signature of interviewer: Date:

A: Type of household

1. HIV/AIDS 2. Non-HIV/AIDS 3. None of the two

RESPONDENT AGREES TO BE INTERVIEWED...................................

RESPONDENT DOES NOT AGREE TO BE INTERVIEWED.......................

RECORD NUMBER

SECTION A. PARTICIPANT IDENTIFICATION

A1: Sex 1. Male 2. Female

A2: Age ……………………………………………….

A3: Present Marital Status 1. Married 2. Consensual Union

3. Separated 4. Divorced 5. Widowed 6. Never Married

A4: Relationship to head of household

1. Head 2. Spouse (Wife/Husband) 3. Child (Son/Daughter)
4. Grandchild 5. Parent/Parent-in-law 6. Son/Daughter-in-law 7. Other Relative 8. Adopter/Foster/Step child 9. House help
10. Non-relative

A5: Occupation

1. Professional/Technical 2. Administrative/Managerial

3. Clerical 4. Sales 5. Service 6. Agric/Anim. Husb./Fishing
7. Pdn & related wks. 8. Home maker

9. Other (Specify) ………………………………………….

A6: Religion

1. Catholic 2. Anglican 3. Presbyterian 4. Methodist

5. Pentecostal/Charismatic 6. Spiritualist 7. Other Christian
8. Muslim 9. Traditional 10. No Religion 11. Other

A7: Ethnicity

1. Akan 2. Ewe 3. Ga 4. Northerner 5. Other

A8: Highest level of education completed

1. None 2. Primary 3. Middle/JHS/JSS 4. Voc./Comm.
5.'O'Level/A'Level/SHS/SSS 6. Post-Secondary 7. Koranic

8. Other (Specify)

A9: Has there been any recent[1] death in this household?

1. Yes 2. No

A10: If Yes to Q1, was it HIV/AIDS-related?

1. Yes 2. No

A11: If No to Q3, can you tell me what the cause was? ------------------------------

A12: Has there been any recent sickness/ailment in this household?

1. Yes 2. No

A13: If Yes to A123, was it HIV/AIDS-related?

1. Yes 2. No

A14: If No to A12 could you tell me what sickness/ailment it was? ---------------

A15: What is your HIV status? 1. Positive 2. Negative 3. Don't Know

A16: If positive to A15, cause of infection

1. Intravenous drug user 2. Sexual intercourse (hetero)

3. Perinatal 4. No Response

[1] Recent means within the past three years

SECTION B: ECONOMIC CHARACTERISTICS OF SURVEYED HOUSEHOLDS

B1: What is your Household size? ………………………………………

B2: What is your average Monthly Household Income …………………………

(**Interviewer to help respondent determine average amount in GH¢**)

B3: Which of the following Assets are owned by your household?

1a. House 1b. How many houses owned ……………………

2a. Farm Land 2b. Size of Farm Land owned (in acres)…………

3a. Building Land 3b. Size of building land (in plots) ………………

4a. Car 4b. Number of cars owned ………………………

5a. Livestock owned 5b. Specify type & number ………………………

B4: Total value (in GH¢) of assets owned ………………………………………

B5: Self-described economic status in community

1. Very well off
2. Well off
3. Moderately well off
4. Poor
5. Very Poor

SECTION C: THE DEPENDENT POPULATION

<u>Before the on-set of disease</u>

C1: Any children < 15 years in the household? 1. Yes 2. No

C1.1: How many of them were you taking care of in terms of <u>Education</u>? ....

C1.2: How many of them were you taking care of in terms of Health?

....

C1.3: How many of them were you taking care of in terms of Upkeep?

.... C2: How much was averagely spent on them yearly on the following?

	List	Annual Expenditure on		
		Education	**Health**	**Upkeep**
C2.1				
C2.2				
C2.3				
C2.4				

C3: Adults in the household

				If **YES**, how many?
C3.1	Any Adults >59 years & less than 65 years old in the household?	1. Yes	2. No	
C3.2	Any Adults >65 years old in the household?	1. Yes	2. No	

C4: How much was averagely spent on them yearly?

	List	Annual Expenditure on	
		Health	**Upkeep**
C4.1			
C4.2			

After the on-set of Disease/Death

C5: How much is averagely spent on the children yearly?

	List	Annual Expenditure on		
		Education	**Health**	**Upkeep**
C5.1				
C5.2				
C5.3				
C5.4				

C6: How much is averagely spent on the adults yearly?

	List	Annual Expenditure on	
		Health	**Upkeep**
C6.1			
C6.2			

C7: How are they being taken care of currently?

C71. The Children………………………………………………………

C72. The Adults…………………………………………………………

SECTION D: DIRECT AND INDIRECT COSTS OF DEATH (HIV/NON-NIV)

I: DIRECT COSTS

D1: Cost of Medical Treatment (Monthly in GH¢) ……………………………

D2: Travel Expenses (For medical treatment in GH¢) …………………………

D3: Funeral Expenses (in GH¢ If person died) …………………………………

II: INDIRECT COSTS

D4: Monthly Income loss of care provider(s) (Monthly in GH¢) ………………

D5: Monthly Income loss of the deceased (Regular job in GH¢) ………………

D6: Monthly Income loss of the deceased (Supplementary job in GH¢) ………

SECTION E: SOCIO-ECONOMIC IMPACTS OF DEATH ON HOUSEHOLDS

E1: Family Labour Supply and Family Production

E1.1. Do you have a Family Business? 1. Yes 2. No

E1.2. Any Impact of the sickness & death on your Family Business? 1. Yes 2. No

E1.3. If Yes to E1.2, nature of impact ………………………………………………………………………………

E2: Impact on children

		State number here
E2.1	Number of affected young children	
E2.2	Number of children being cared for by a parent	
E2.3	Number of children being cared for by a grandparent	
E2.4	Number of children being cared for by other relations	
E2.5	Number of children being cared for by orphanage, church, etc	

E3: Impact on the Elderly

E3.1 Is/Was there an elderly in this household who was taking care of the deceased before death? 1. Yes 2. No

E3.2 If the response to E3.1 is Yes, how many of them were/are aged >60 yrs? … …………………

E3.3 Number of elderly being cared for by the deceased before death ………………………………..

E3.4 Does household have elderly looking after themselves?
1. Yes 2. No

E3.5 If the response to E3.4 is Yes, how many ………………………………………………………..

E3.6 How many elderly are being cared for by spouse or children of the deceased? …………………

E3.7 How many are being cared for by community, church, etc? ……………………………………

E4: Has there been any form of Social Discrimination against infected/affected persons in this household?

1. Discriminated against
2. Not discriminated against
3. No response

E5: If response to E4 is (1), what was the form of social discrimination? (**Multiple response acceptable**)

1. Household member forced to leave job

2. No customers
3. Departing employees
4. No goods orders for family business
5. No association
6. Forced to leave community
7. Children forced to leave school
8. Children prevented from playing with other children
9. Others (Specify) ………………………………………………

<u>Household debt</u>

E6: Does household have any debt as a result of the sickness/death?

1. Yes 2. No

E7: If Yes to E6, how did the debt come about?

……………………………………………………………………………

……………………………………………………………………………

<u>Household Coping Strategy during illness and after Death</u>

E7: Dissaving as a result of illness/death

E7.1 Is household using up savings as a result of the illness/death?

1. Yes 2. No

E7.2 If Yes, what is the average savings used per month (in GH¢)

………………………………...

E8: Consumption expenditure reduction as a result of illness/death

E8.1 Is household reducing expenditure? 1. Yes 2. No

E8.2 Any change in household food consumption?

1. Yes 2. No

E8.3 If yes to E8.2, please indicate this change (in percent)

…………………………………….

E9: Sale of household assets in order to take care of patient

E9.1 Is household selling assets in order to take care of patient? 1.Yes 2. No

E9.2 If Yes to E9.1, specify (**Multiple response accepted**)

1. Land 2. Livestock 3. vehicle 4. other (specify)

E10: Was there a reallocation of household member's time as a result of the sickness/death? [**Multiple responses accepted**]

1. No change
2. Worked harder to substitute for lost income
3. Needed to find job
4. Helped with family business
5. Left job to help take care of the sick person
6. Reduced work time to help family
7. Changed to new job for higher income
8. Found supplementary job
9. Left school for work
10. Others (Specify) ……………………………………………………

E11.1: If response for E10 is (9), then Sex of child who left school for work

1. Male 2. Female

E11.2: If response for E10 is (9), then Age of child who left school for work ....

E12: How was lost labour in family production substituted?

1. Employed substitute labour
2. Other members worked harder since there was no substitution
3. Member(s) left school for family work
4. No response/Not appropriate

E13: Borrowing

E13.1 Did household borrow as a result of sickness/death of a member?

1. Yes 2. No

E13.2 If Yes, how? Amount

1. From a bank

………………………………..

2. From money lender …………………………………

3. From relatives …………………………………

4. Cooperatives/revolving funds ………………………………...

E14: Transfers-in

E14.1 Has your family received any transfers-in since the onset of disease or death? 1. Yes 2. No

E14.2 If Yes, amount per month (GH¢) ……………………………..

E15: Non-family institutions and health care costs

E15.1 Did family receive government health care benefits for government employees? 1. Yes 2. No

E15.2 Did family receive benefits from National Health Insurance Scheme? 1. Yes 2. No

E15.3 Did family receive Health care benefits for PLWHAs?

1. Yes 2. No

E15.4 Did family receive benefits from any social security Programme?

1. Yes 2. No

E15.5 Specify other benefits received …………………………………

Households Experiencing HIV/AIDS Death

E16: What is currently the appropriate impact of death on household consumption?

1. Very serious impact on consumption
2. Moderate impact on consumption
3. No impact on consumption

E17 What is currently the source of household expense for health care?

1. Household savings
2. Selling assets
3. Borrowing
4. Others (Specify) …………………………………………

E18 What is currently the time re-allocation of Children?

1. Had to find job
2. Left school for work
3. Left school to look after siblings, children, sick person, household chores, etc.
4. Others (Specify) …………………………………………

E19.1 What is the sex of the child who left school for work?

1. Male 2. Female

E19.2 What is the age of the child who left school for work?

..................….

E20 Who currently cares for the children?

1. Under care of extended family
2. Under care of orphanage
3. Child taking care of self
4. Others (Specify) …………………………………………

E21 Who currently cares for the Elderly?

1. Elderly looking after themselves
2. Elderly under care of extended family
3. Elderly under care of community, NGO, etc.
4. Others (Specify) …………………………………………

<u>Household with non-HIV/AIDS- related death</u>

E22: What is currently the appropriate impact of death on household consumption?

1. Household feels very serious impact on consumption
2. Household feels moderate impact on consumption
3. Household feels no impact on consumption

E23: What is currently the source of household expense for health care?

1. Household savings
2. Selling assets
3. Borrowing
4. Others (Specify) …………………………………………

E24: What is currently the time re-allocation of Children?

1. Had to find job
2. Left school for work
3. Left school to look after siblings, children, sick person, house chores, etc.
4. Others (Specify) …………………………………………

E25.1: What is the sex of the child who left school for work?

1. Male 2. Female

E25.2: What is the age of the child who left school for work?

………….

E26 Who currently cares for the children?

1. Under care of extended family
2. Under care of orphanage
3. Child taking care of self
4. Others (Specify) …………………………………………

E27 Care of Elderly after death

1. Elderly looking after themselves
2. Elderly under care of extended family
3. Elderly under care of community, NGO, etc.
4. Others (Specify) …………………………………………

SECTION F: DATA ON DECEASED

F1: What was the Age of the deceased at the time of death?

………………………………………………..

F2: What was the Sex of the deceased? 1. Male 2. Female

F3: What was the Marital Status of the deceased?

1. Married 2. Consensual Union 3. Separated 4. Divorced 5. Widowed 6. Never Married

F4: What was the Household status of the deceased?

1. Head 2. Spouse (Wife/Husband) 3. Child (Son/Daughter)

4. Grandchild 5. Parent/Parent-in-law 6. Son/Daughter-in-law 7. Other Relative 8. Adopter/Foster/Step child 9. House help 10. Non-relative

F5: What was the cause of death?

………………………………………………………………………

F6: What was the Occupation of the deceased before death?

1. Professional/Technical 2. Administrative/Managerial 3. Clerical

4. Sales 5. Service 6. Agric/Anim. Husb./Fishing

7. Pdn & related wk 8. Home maker 9. Other (Specify) …………

F7: What was the Income of the deceased at the time of death? ………………

F8: What was the Educational attainment of the deceased?

1. None 2. Primary 3. Middle/JHS 4. Voc./Comm.
5. 'O' Level 6. SHS 7. 'A' Level 8. Training College

9. Techn./Prof. 10. Tertiary 11. Koranic 12. Don't Know

F9: What was the Year of death ………………………………….

F10: What was the Cause or symptom(s) the deceased showed before death

…………………………………………………………………….

F11 For how long was the sickness suffered by the deceased (In months)

………….

F12: What immediate action was taken about the illness?

…………………………………………………………………….

F13: Where did the person die?

1. At home 2. Hospital 3. Prayer Camp
4. Other (Specify) …………………………..

Appendix C

Sample SAS Codes Used in the Modelling Process

The General SAS code written by this researcher and used was of the form:

proc glimmix

data=[*specify data-file path*] ic=pq method= [*specify the method here*];

class [*list variables to be used in the model (both fixed effects and random effects variables*)];

model [*specify the model to be used*] / dist=binary link=logit solution ddfm=contain;

random [*specify the g-side random effect(s)*]/ solution;

random _residual_ / subject=intercept type=ar(**1**);

run;

SAS code A

A sample of the specific SAS codes which were used in the modelling process is presented below:

```
proc  glimmix
    data='c:\users\david\desktop\phd_thesis_last_lap_26122013\david_final_phd_data_20072013.sas7bdat' ic=pq method= rmpl;
      class   a3 agegroup a5maristat a9ethnicity a10edn a14r type;
      model e101_rec =a3 agegroup a5maristat a10edn a14r / dist=binary link=logit solution ddfm=contain;
      random type/ solution;
      random _residual_ / subject=intercept type=ar(1);
    run;
```

Appendix D

Sample SAS Output

Appendix D1: Unedited output of PROC GLIMMIX on Maximum Pseudo Likelihood

Appendix D1.1 Model Information

Model Information	
Data Set	TMP5.DAVID_FINAL_PHD_DATA_26122013
Response Variable	E101_Rec
Response Distribution	Binary
Link Function	Logit
Variance Function	Default
Variance Matrix	Not blocked
Estimation Technique	MPL
Degrees of Freedom Method	Containment

Appendix D1.2: Class Level Information

Class Level Information		
Class	**Levels**	**Values**
A3	2	0 1
AgeGroup	3	0 1 2
A5MariStat	2	0 1
A10Edn	3	1 2 3
A11	2	0 1
TYPE	3	1 2 3

Appendix D1.3: Number of observations used

Number of Observations Read	601
Number of Observations Used	601

Appendix D1.4: Response Profile

Response Profile		
Ordered Value	**E101_Rec**	**Total Frequency**
1	0	545
2	1	56
The GLIMMIX procedure is modeling the probability that E101_Rec='0'.		

Appendix D1.5: Dimensions

Dimensions	
G-side Cov. Parameters	1
R-side Cov. Parameters	2
Columns in X	13
Columns in Z	3
Subjects (Blocks in V)	1
Max Obs per Subject	601

Appendix D1.6: Optimization Information

Optimization Information	
Optimization Technique	Newton-Raphson with Ridging
Parameters in Optimization	2
Lower Boundaries	2
Upper Boundaries	1
Fixed Effects	Profiled
Residual Variance	Profiled
Starting From	Data

Appendix D1.7: Iteration History

Iteration History					
Iteration	**Restarts**	**Subiterations**	**Objective Function**	**Change**	**Max Gradient**
0	**0**	**5**	2656.5487547	1.22080297	1.07E-7
1	**0**	**3**	2949.0652239	0.56516252	5.058E-6
2	**0**	**3**	3066.7390251	0.10299464	6.659E-8
3	**0**	**2**	3082.3264546	0.04000977	1.132E-7
4	**0**	**1**	3082.631369	0.00794225	2.003E-6
5	**0**	**1**	3084.0858988	0.00426514	1.556E-6
6	**0**	**1**	3084.374265	0.00099333	1.644E-7
7	**0**	**1**	3084.5129855	0.00049828	1.049E-8
8	**0**	**1**	3084.5448834	0.00011025	2.21E-9
9	**0**	**1**	3084.560036	0.00006109	9.42E-11
10	**0**	**1**	3084.5632467	0.00001267	3.08E-11
11	**0**	**1**	3084.5649355	0.00000799	3.746E-7
12	**0**	**1**	3084.565251	0.00000154	1.15E-12
13	**0**	**0**	3084.5654454	0.00000114	3.339E-6

Iteration History					
Iteration	**Restarts**	**Subiterations**	**Objective Function**	**Change**	**Max Gradient**
14	**0**	**1**	3084.5654712	0.00000028	8.466E-9
15	**0**	**0**	3084.5654938	0.00000017	1.57E-7
16	**0**	**0**	3084.5654956	0.00000004	9.87E-7
17	**0**	**0**	3084.5654984	0.00000002	7.813E-7
18	**0**	**0**	3084.5654986	0.00000001	9.589E-7

Convergence criterion (PCONV=1.11022E-8) satisfied.

Appendix D1.8: Fit Statistics

Fit Statistics	
-2 Log Pseudo-Likelihood	3084.57
Pseudo-AIC	3106.57
Pseudo-AICC	3107.01
Pseudo-BIC	3096.65
Pseudo-CAIC	3107.65
Pseudo-HQIC	3086.63
Generalized Chi-Square	434.47
Gener. Chi-Square / DF	0.72
Fit statistics based on pseudo-likelihoods are not useful for comparing models that differ in their pseudo-data.	

Appendix D1.9: Covariance Parameter Estimates

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Estimate**	**Standard Error**
TYPE		2.1252	1.7817
AR(1)	Intercept	-0.03521	0.04195
Residual		0.7229	0.04185

Appendix D1.10: Solutions for Fixed Effects

Solutions for Fixed Effects									
Effect	**Sex**	**Age groups**	**Current Marital Status**	**A10Edn**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept					1.5110	1.2288	2	1.23	0.3439
A3	0				0.3572	0.2570	591	1.39	0.1651
A3	1				0	.	.	.	.
AgeGroup		0			0.2009	0.4540	591	0.44	0.6583
AgeGroup		1			-0.6476	0.4261	591	-1.52	0.1291
AgeGroup		2			0	.	.	.	.
A5MariStat			0		-0.2873	0.2829	591	-1.02	0.3103
A5MariStat			1		0	.	.	.	.
A10Edn				1	0.7033	0.8317	591	0.85	0.3981
A10Edn				2	0.5488	0.8289	591	0.66	0.5082
A10Edn				3	0	.	.	.	.
A14R					0.8181	0.3400	591	2.41	0.0164

Appendix D1.11: Type III Tests of Fixed Effects

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
A3	1	591	1.93	0.1651
AgeGroup	2	591	3.92	0.0204
A5MariStat	1	591	1.03	0.3103
A10Edn	2	591	0.45	0.6348
A14R	1	591	5.79	0.0164

Appendix D1.12: Solution for Random Effects

Solution for Random Effects						
Effect	**Type of household**	**Estimate**	**Std Err Pred**	**DF**	**t Value**	**Pr > |t|**
TYPE	1	-2.0377	0.8638	591	-2.36	0.0187
TYPE	2	0.9183	0.8724	591	1.05	0.2929
TYPE	3	1.1194	0.8645	591	1.29	0.1959

Appendix D2

Unedited output of PROC GLIMMIX on Residual Pseudo Likelihood

Appendix D2.1: Model Information

Model Information	
Data Set	TMP5.DAVID_FINAL_PHD_DATA_20072013
Response Variable	E101_Rec
Response Distribution	Binary
Link Function	Logit
Variance Function	Default
Variance Matrix	Not blocked
Estimation Technique	Residual MPL
Degrees of Freedom Method	Containment

Appendix D2.2: Class Level Information

Class Level Information		
Class	**Levels**	**Values**
A3	2	0 1
AgeGroup	3	0 1 2
A5MariStat	2	0 1
A9Ethnicity	2	1 2
A10Edn	3	1 2 3
A14R	2	0 1
TYPE	3	1 2 3

Appendix D2.3: Number of observations

Number of Observations Read	601
Number of Observations Used	601

Appendix D2.4: Response Profile

Response Profile		
Ordered Value	**E101_Rec**	**Total Frequency**
1	0	545
2	1	56
The GLIMMIX procedure is modeling the probability that E101_Rec='0'.		

Appendix D2.5: Dimensions

Dimensions	
G-side Cov. Parameters	1
R-side Cov. Parameters	2
Columns in X	13
Columns in Z	3
Subjects (Blocks in V)	1
Max Obs per Subject	601

Appendix D2.6: Optimization Information

Optimization Information	
Optimization Technique	Newton-Raphson with Ridging
Parameters in Optimization	2
Lower Boundaries	2
Upper Boundaries	1
Fixed Effects	Profiled
Residual Variance	Profiled
Starting From	Data

Appendix D2.7: Iteration History

Iteration History					
Iteration	**Restarts**	**Sub-iterations**	**Objective Function**	**Change**	**Max Gradient**
0	**0**	**5**	2662.1708069	1.26719169	5.845E-8
1	**0**	**3**	2950.5914766	0.58589711	2.586E-6
2	**0**	**3**	3065.2601882	0.10779947	3.174E-8
3	**0**	**2**	3079.7167923	0.04092001	3.807E-8
4	**0**	**1**	3079.8025312	0.00783510	1.885E-6
5	**0**	**1**	3081.2125041	0.00439408	7.214E-7
6	**0**	**1**	3081.4730724	0.00099456	1.153E-7
7	**0**	**1**	3081.6075175	0.00051607	6.369E-9
8	**0**	**1**	3081.6367003	0.00011068	1.568E-9
9	**0**	**1**	3081.6515237	0.00006338	5.91E-11
10	**0**	**1**	3081.6544331	0.00001268	2.09E-11
11	**0**	**1**	3081.656113	0.00000819	6.56E-13
12	**0**	**1**	3081.6563915	0.00000152	6.81E-13

Iteration History					
Iteration	**Restarts**	**Sub-iterations**	**Objective Function**	**Change**	**Max Gradient**
13	**0**	**0**	3081.6565845	0.00000115	4.263E-6
14	**0**	**1**	3081.6566098	0.00000030	1.71E-12
15	**0**	**0**	3081.6566325	0.00000018	1.315E-7
16	**0**	**0**	3081.6566339	0.00000003	1.154E-6
17	**0**	**0**	3081.656637	0.00000002	1.001E-6
18	**0**	**0**	3081.6566371	0.00000001	1.156E-6

Convergence criterion (PCONV=1.11022E-8) satisfied.

Appendix D2.8: Fit Statistics

Fit Statistics	
-2 Res Log Pseudo-Likelihood	3081.66
Pseudo-AIC	3103.66
Pseudo-AICC	3104.11
Pseudo-BIC	3151.89
Pseudo-CAIC	3162.89
Pseudo-HQIC	3122.44
Generalized Chi-Square	433.22

Fit Statistics	
Gener. Chi-Square / DF	0.73
REML information criteria are adjusted for fixed effects and covariance parameters. Fit statistics based on pseudo-likelihoods are not useful for comparing models that differ in their pseudo-data.	

Appendix D2.9: Covariance Parameter Estimates

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Estimate**	**Standard Error**
TYPE		3.2372	3.2948
AR(1)	Intercept	-0.03362	0.04195
Residual		0.7306	0.04253

Appendix D2.10: Solutions for Fixed Effects

Solutions for Fixed Effects										
Effect	**Sex**	**Age groups**	**Current Marital Status**	**A10Edn**	**Has there been any recent sickness or ailment in this household?**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept						2.3268	1.3848	2	1.68	0.2349
A3	0					0.3623	0.2580	591	1.40	0.1609
A3	1					0	.	.	.	.
AgeGroup		0				0.1905	0.4564	591	0.42	0.6765
AgeGroup		1				-0.6475	0.4286	591	-1.51	0.1313
AgeGroup		2				0	.	.	.	.
A5MariStat			0			-0.2860	0.2841	591	-1.01	0.3145
A5MariStat			1			0	.	.	.	.
A10Edn				1		0.7038	0.8363	591	0.84	0.4004
A10Edn				2		0.5431	0.8334	591	0.65	0.5149
A10Edn				3		0	.	.	.	.
A14R					0	-0.8138	0.3420	591	-2.38	0.0177
A14R					1	0	.	.	.	.

Appendix D2.11: Type III Tests of Fixed Effects

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
A3	1	591	1.97	0.1609
AgeGroup	2	591	3.80	0.0229
A5MariStat	1	591	1.01	0.3145
A10Edn	2	591	0.46	0.6310
A14R	1	591	5.66	0.0177

Appendix D2.12: Solution for Random Effects

Solution for Random Effects						
Effect	**Type of household**	**Estimate**	**Std Err Pred**	**DF**	**t Value**	**Pr > |t|**
TYPE	1	-2.0531	1.0572	591	-1.94	0.0526
TYPE	2	0.9271	1.0643	591	0.87	0.3841
TYPE	3	1.1260	1.0577	591	1.06	0.2875

Appendix E

Proofs

$$L(\boldsymbol{\beta}) = (2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right)$$

Proof:

For a set of n independently and identically distributed (IID) observations, y_1, y_2, ... y_n, the likelihood function can take the form

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{n} f(y_i;\ \mu, \sigma^2)$$

$$= \prod_{i=1}^{n}(2\pi\sigma^2)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\frac{(y_i-\mu)^2}{\sigma^2}\right)$$

$$= (2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right)$$

$$l(\boldsymbol{\beta}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2$$

Normal Distribution:

$$l(\boldsymbol{\beta}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2$$

Proof:

Taking the natural logarithm of the likelihood function $L(\boldsymbol{\beta})$ for a set of n independently and identically distributed (IID) observations, $y_1, y_2, \ldots y_n$ with the normal distribution, we have

$$l(\mu, \sigma^2;\ y_1, y_2, \ldots\ y_n) = \ln[L(\mu, \sigma^2;\ y_1, y_2, \ldots\ y_n)]$$

$$= \ln\left((2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2 \right) \right)$$

$$= \ln\left((2\pi\sigma^2)^{-\frac{n}{2}} \right) + \ln\left(\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2 \right) \right)$$

$$= -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2$$

$$= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2$$

Poisson Distribution:

$$l(\boldsymbol{\mu}; \boldsymbol{Y}) = \sum_{i=1}^{n} y_i \ln \mu - n\mu$$

Proof:

Similar to the normal distribution above, for a set of n independently and identically distributed (IID) observations, $\boldsymbol{Y} = y_1, y_2, \ldots y_n$ with the Poisson distribution and a parameter μ, we have

$$L(\boldsymbol{\mu};\boldsymbol{Y}) = \prod_{i=1}^{n} \frac{\mu^{y_i} e^{-\mu}}{y_i!}$$

$$= \frac{\mu^{\sum_{i=1}^{n}} e^{-n\mu}}{y_1!\ y_2! \ldots y_{n!}}$$

Now, taking the natural logarithm of the likelihood function $L(\boldsymbol{\mu};\boldsymbol{Y})$, we have

$$l(\boldsymbol{\mu};\boldsymbol{Y}) = \ln L(\boldsymbol{\mu};\boldsymbol{Y}) = \ln\left(\frac{\mu^{\sum_{i=1}^{n}} e^{-n\mu}}{y_1!\ y_2! \ldots y_{n!}}\right)$$

$$= ln\left(\mu^{\sum_{i=1}^{n} y_i e^{-n\mu}}\right)$$

$$= \sum_{i=1}^{n} y_i \ln \mu - n\mu$$

Appendix F

Components of GLMs

Table F1.1: Components of GLMs

$Y \sim$	Normal (u, σ^2)	Gamma (α, β)	Poisson (λ)	Bin. $(m, q)/m$
Link g	Identity	Reciprocal	Log	Logit
$\mathbb{E}(Y) = \mu(\theta)$	$\theta = \mu$	$-\theta^{-1} = \frac{\alpha}{\beta}$	$e^{\theta} = \lambda$	$\frac{e^{\theta}}{1+e^{\theta}} = q$
$\mathbb{V}(Y) = V(\mu)\emptyset$	σ^2	$\frac{1}{\theta^2\ \alpha} = \frac{\alpha}{\beta^2}$	$e^{\theta} = \lambda$	$\frac{q(1-q)}{m}$
$V(\mu)$	1	θ^{-2}	$e^{\theta} = \lambda$	$q(1-q)$
ϕ	σ^2	α^{-1}	1	$1/m$
$c(y, \phi)$	$-\frac{1}{2}\left[\frac{y^2}{\sigma^2} + \ln(2\pi\sigma^2)\right]$	α In αy +In y -In $\Gamma(\alpha)$	-In $(y!)$	$\ln\binom{m}{my}$