

**UNIVERSITY OF GHANA**

**USING GENERALIZED ESTIMATING EQUATIONS TO ASSESS THE EFFECT OF  
THE SCHOOL FEEDING PROGRAMME AND THE CAPITATION GRANT ON BASIC  
SCHOOL ENROLMENT. A CASE STUDY OF THE WENCHI MUNICIPALITY.**

**BY**  
**DANIEL OKYERE**  
**(10442620)**

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

**INTEGRI PROCEDAMUS**

**JULY, 2015.**

## DECLARATION

### Candidate's Declaration

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

SIGNATURE: ..... DATE.....

**DANIEL OKYERE**

**(10442620)**

### Supervisors' Declaration

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

SIGNATURE: ..... DATE.....

**DR SAMUEL IDDI**

(Principal Supervisor)

SIGNATURE: ..... DATE.....

**DR ISAAC BAIDOO**

(Co-Supervisor)

## ABSTRACT

Education has been defined as the act or process of imparting or acquiring general knowledge, developing the process of reasoning and judgment, and generally of preparing oneself or others intellectually for mature life. It is an unarguable fact across the globe that a nation's development rests on the education of its citizens. But for a country to achieve a high literacy rate, basic education, which is regarded as the foundation should be held in high esteem. For this reason, governments in the pre and post-colonial era made significant strides all in an attempt to universalize primary education. These attempts brought in its wake a lot of educational and social interventions, policies and initiatives. In recent times in Ghana, two of such interventions believed to address the problem of universalizing primary education are the school feeding programme and the capitation grant, introduced in 2005 and 2004 respectively. In view of this, the study sought to assess the marginal effect of these two interventions on basic school enrolment using the Wenchi municipality as a case study. All eighty-one (81) primary schools (both public and private) in the municipality were used for the study. Enrolment figures of schools together with their feeding and capitation grant status, covering the period 2008 to 2015 were obtained at the Education Management Information Systems (EMIS) department of education directorate. Other variables obtained in addition were; the number of trained/untrained teachers, number of classrooms in good condition, number of textbooks per child, location of school and availability of portable water and toilet facilities. The data obtained was organized using MS-Excel and subsequently analyzed using Generalized Estimating Equations (GEE) family of models. The analysis was done using R and SPSS. The findings from the analyses revealed that the feeding programme and the capitation grant contribute significantly in increasing enrolment. However, other identifiable factors which were found to affect enrolment were number of trained teachers and availability of toilet facilities.

## DEDICATION

I dedicate this thesis to my beloved parents; Mr. Samuel Okyere and Miss. Constance Afful, and also to my brothers; Enock Danso Okyere and Frederick Hoppey.



## ACKNOWLEDGEMENT

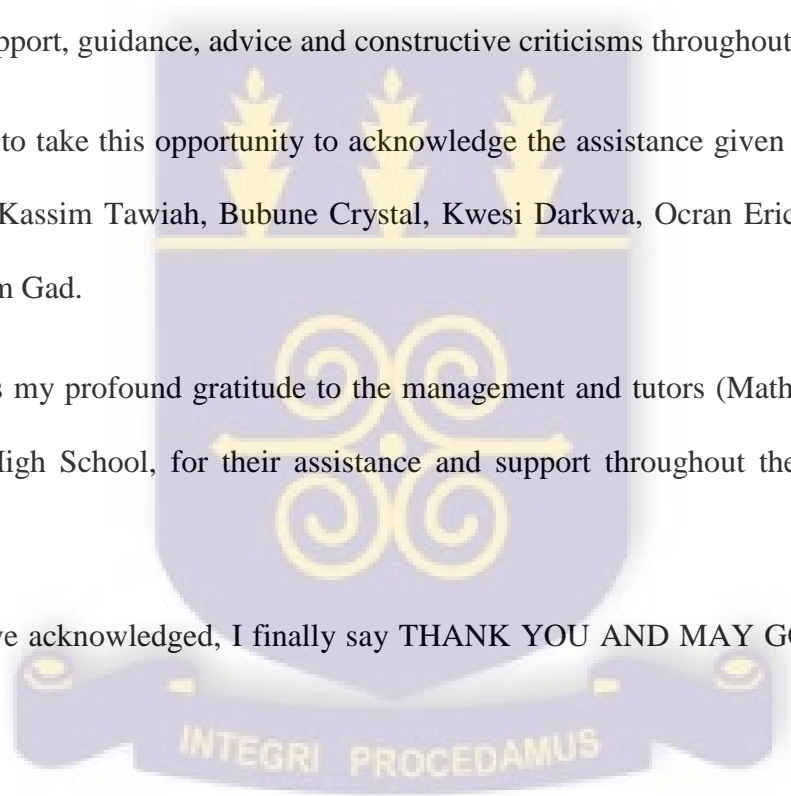
To God be the glory, great things he has done. I will first and foremost give thanks to the Almighty God for giving me the strength, guidance and knowledge throughout my period of study. If not for his mercies and abundant blessings, I wouldn't have pursued education up to this level.

Next, I render my heart- felt thanks to my supervisors; Dr. Samuel Iddi and Dr. Isaac Baidoo for their countless support, guidance, advice and constructive criticisms throughout my study.

I would also like to take this opportunity to acknowledge the assistance given me by my course mates especially Kassim Tawiah, Bubune Crystal, Kwesi Darkwa, Ocran Eric, Sarkodie Curtis and Kusi- Boadum Gad.

I wish to express my profound gratitude to the management and tutors (Mathematics Dept.) of Wenchi Senior High School, for their assistance and support throughout the entire period of study.

To all those I have acknowledged, I finally say THANK YOU AND MAY GOD BLESS YOU ALL.



## TABLE OF CONTENTS

### Contents

Declaration.....	i
Candidate’s Declaration.....	i
Supervisor’s Declaration.....	i
Abstract .....	ii
Dedication.....	iii
Acknowledgement.....	iv
Table of Contents.....	v
List of Figures.....	ix
List of Tables.....	ix

### CHAPTER ONE: INTRODUCTION

1.0 Background of the Study.....	1
1.1 Statement of the Problem.....	4
1.2 Objectives of the Study.....	5
1.3 The Scope of the Study.....	5
1.4 Data Description.....	5
1.5 The Study Area.....	6
1.6 Data Source and Organization.....	7
1.7 Data Analyses.....	7
1.8 Significance of the Study.....	8
1.9 Organization of the Study.....	8

**CHAPTER TWO: LITERATURE REVIEW**

2.0 Introduction.....9

2.1 Feeding in Basic Schools.....9

    2.1.1 Brain Functioning in Children.....10

    2.1.2 Basic School Enrolment, Attention and Retention.....10

2.2 Ghana’s School Feeding Programme.....11

2.3 School Fees Abolition and the Capitation Grant.....12

2.4 Evaluation of Educational Interventions.....14

2.5 Modeling Longitudinal Data.....16

2.6 Marginal Modeling.....17

2.7 Generalized Estimating Equation.....18

**CHAPTER THREE: METHODOLOGY**

3.0 Introduction.....20

3.1 The Exponential Family of Distributions.....25

    3.1.1 Introduction.....25

    3.1.2 Properties of the Exponential Family.....26

3.2 Generalized Linear Models.....29

    3.2.1 Introduction.....29

    3.2.2 GLM for Count Data.....31

3.3 Estimation Methods for GLM’s.....33

    3.3.1 Maximum Likelihood Estimation.....33

    3.3.2 Quasi-Likelihood Estimation.....36

    3.3.3 Parameter Estimation for the Poisson Regression Model.....37

3.4 Significance of Regression Parameters.....39

3.5 Models for Correlated Data.....	41
3.5.1 The Conditional Model.....	41
3.5.2 The Marginal Model.....	43
3.6 The Generalized Estimating Equations.....	44
3.6.1 Introduction.....	44
3.6.2 GEE Method Outline.....	45
3.6.3 Working Correlation Forms.....	46
3.6.4 GEE Estimation.....	50
3.6.5 Inference.....	52
3.6.6 Limitations of the GEE.....	53

**CHAPTER FOUR: DATA ANALYSES**

4.0 Introduction.....	54
4.1 Definition of Variables.....	54
4.2 Preliminary Analyses.....	55
4.3 Further Analyses.....	60
4.3.1 Analyses of GEE Family of Models.....	62
4.3.2 GEE Independence Model.....	62
4.3.3 GEE Exchangeable Model.....	64
4.3.4 GEE AR-1 Model.....	65
4.3.5 GEE Unstructured Model.....	66
4.4 Inference.....	68

**CHAPTER FIVE: DISCUSSION, CONCLUSION AND RECOMMENDATIONS**

5.0 Introduction.....69

5.1 Discussion.....69

5.2 Conclusion.....70

5.3 Recommendations.....71

References.....73

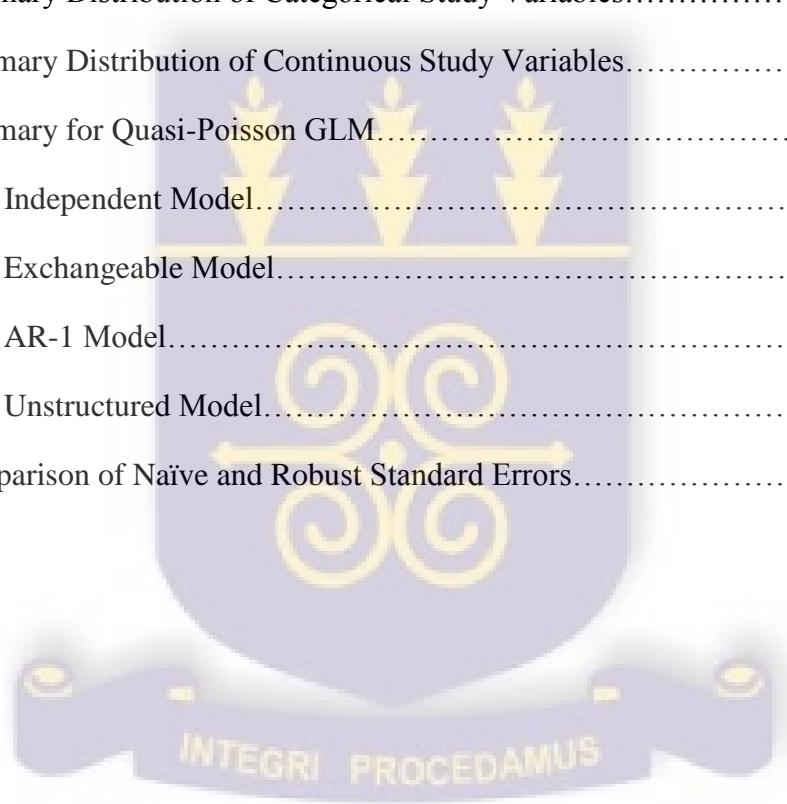


## LIST OF FIGURES

Figure 4.1: Plot of Total Enrolment by Location of School.....	57
Figure 4.2: Mean Plot for School Feeding.....	58
Figure 4.3: Mean Plot for Capitation Grant.....	59

## LIST OF TABLES

Table 4.2.1: Summary Distribution of Categorical Study Variables.....	55
Table 4.2.2: Summary Distribution of Continuous Study Variables.....	56
Table 4.3.1: Summary for Quasi-Poisson GLM.....	61
Table 4.3.2: GEE Independent Model.....	63
Table 4.3.3: GEE Exchangeable Model.....	64
Table 4.3.4: GEE AR-1 Model.....	65
Table 4.3.5: GEE Unstructured Model.....	66
Table 4.3.6: Comparison of Naïve and Robust Standard Errors.....	67



## **CHAPTER ONE**

### **INTRODUCTION**

#### **1.0 Background of the Study**

Education is defined as the act or process of imparting or acquiring general knowledge, developing the process of reasoning and judgment, and generally of preparing oneself or others intellectually for mature life. It is in the light of this that John Dewey (1938) defined education as the continuous reconstruction or reorganization of experience which adds to the meaning of life.

Every nation's or community's development to a very large extent is dependent on the quality of education enjoyed by its citizens, since it is generally believed that the bedrock for any true development must begin with the development of its human resources. That is, there is general consensus across the globe that one of the most important means by which poverty and inequality can be effectively addressed is through human development and the principles of human development are built on sound education. Quality education might help individuals to grow, develop, earn a decent living in the society and contribute positively to the welfare of the society in which they live.

Apart from the direct benefits of building skills, increasing knowledge and opening for critical thinking, quality education could also be considered to be a reliable tool for poverty reduction especially in sub-Saharan Africa.

For quality education to result in the development of a community, the basic level, which is the foundation, must be held in high esteem. It is at this stage that one gains fundamental knowledge, values and attitudes required for the full development of one's self in order to ensure effective participation in the socio-economic development of the community in which one lives.

For the young ones to acquire and develop the skills, knowledge, values and attitudes that will help them to contribute positively to the social, economic, political and moral upliftment and advancement of the society in which he or she lives (and the nation at large), education at the basic level must be of much quality.

Pre- colonial and post- colonial governments over the years have made significant strides all in an attempt to increase access and participation, ensure gender equity and improve quality of education in basic schools, and this led to the introduction and implementation of various educational reforms, initiatives, intervention and policies. The current basic education structure and curriculum has its roots in Ghana's colonial past. Pre-independence education was characterized by attempts to create incentives for all children to attend school, as happened in Northern Ghana with the introduction of free education to improve access. The earliest sign of a plan to universalize primary education was in 1945 when the colonial government proposed a 10-year plan to universalize primary education in 25 years based on cost projections set within affordable limits. The next, and most significant, was in 1951, when Dr. Kwame Nkrumah implemented the Accelerated Development Plan of education aimed at rapidly improving access into basic (elementary) and secondary education. The policy sought to universalize primary education to meet the high demand for education. However, the 1951 educational policy faced major infrastructural challenges and was therefore reviewed in 1966. The Kwapong review committee was implored to address the problem of the majority of pupils from elementary schools who could not gain entry into the limited number of places into the then "grammar" schools. The Kwapong committee introduced the concept of continuation schools, but which was later criticized as being one which only favours the rich in society.

The major reforms in Ghana's educational history during the post-colonial era began in 1974, following the criticisms of the reforms implemented as a result of the recommendation by the Kwapong committee. The Dzobo review committee of 1974 introduced the concept of comprehensive Junior Secondary Schools to teach academic and practical skills into pupils. This was shortly followed by the education reforms of 1978 which led to the introduction of the junior and secondary school concept.

Education, as being recognized as the fundamental building block of the economy is sacrosanct, and therefore enshrined in the 1992 constitution of Ghana. Article 38 of the 1992 constitution enjoins government to provide access to free compulsory universal basic education (fCUBE) to all children of school going age. The fCUBE policy was fully launched in 1996 to provide access to quality education to all children of school going age. It is in pursuance of this requirement that Ghana subscribed to some international educational initiatives such as Education for All (EPA) and the UN Millennium Development Goals (MDG's).

Despite all these policies, interventions and strategies put in place by previous regimes to ensure universal primary and quality education in basic education, the ultimate was never reached. Research works have shown that pupils supposed to be in school failed to enroll, and even those who are already enrolled drops out primarily due to financial constraints and other factors. Other researchers also believe that the inability of school children to enjoy at least two square meals a day affects nutritional needs necessary for effective teaching and learning. Against this background, the Kuffour administration launched the capitation grant policy in 2004 and subsequently the school feeding program in 2005. These two policies were purposely to increase access, ensure retention, improve quality, and ensure gender equity, among others.

The capitation grant makes basic education free from any form of school fees. Generally, this should ensure universal primary education, high enrolments and retention in basic schools. The school feeding program provides free meals to pupils, thereby providing for their nutritional needs needed to enhance their learning capabilities. It is also enrolment driven and indirectly expected to translate into higher performance by pupils. Due to inadequate funds the school feeding program is piloted and so does not cover all schools.

### **1.1 Statement of the problem**

It is generally and widely accepted across the globe that poverty can be effectively eradicated through quality human resource development, and the principle of quality human development are built on very sound education. The importance of education as a fundamental tool for development is clearly articulated in the United Nation's MDG's and the EPA.

In recognition of the key role of education to a nation's development, the government of Ghana subscribed to the principles of the MDG's and EPA. It is then required by every government, upon subscription, to adopt measures and policies which will help attain universal primary education within a stipulated time frame without compromising quality. In line with goal II of the MDG's, universal primary education is supposed to be achieved by the year 2015.

As part of the strategies to achieve universal primary education, the government of Ghana lunched the capitation grant and school feeding program. These policies are expected to increase enrolments significantly in basic schools.

This study thus seeks to evaluate the impact that the school feeding program and the capitation grant have brought to bear on basic school enrolment.

## **1.2 Objectives of the Study**

The general aim of the study is to use GEE family of models to study the effect of identified variables on enrolment in basic schools.

The specific objectives of the study are;

- (i) To fit GEE family of models under different working correlation assumptions and to choose the most appropriate model for the data.
- (ii) To investigate whether some of the mentioned variables have effect on basic school enrolment.

## **1.3 Scope of the Study**

The study will consider enrolment data from all 81 primary schools (both private and public) in the Wenchi municipality covering a period of eight (8) years, from 2008-2015. Also, records of other covariates, which might tend to affect enrolment in schools are obtained.

## **1.4 Data description**

The research study is to assess the overall effect of the school feeding and the capitation grant on general enrolments in basic schools using the Wenchi municipality as a case study. Enrolment

figures from all the primary schools within the municipality, starting from 2008 to 2015 were obtained. Alongside the enrolment figures are data on other covariates such as the number of trained/untrained teachers, number of text books per child, school feeding, capitation grant, geographical location and status of school, number of classrooms in good condition and availability of portable water and toilet facilities in the schools. The data therefore constitutes a longitudinal study with repeated measures on the same school overtime with several observed covariates.

It is however important to state that the Wenchi municipality was chosen as a case study mainly due to the fact that it was categorized as “deprived district” before the inception of the school feeding program. Secondly, we chose Wenchi as a case study as a result of convenience and easy access to data.

### **1.5 The study area**

The Wenchi Municipal is located in the Western part of Brong Ahafo Region. It is situated at the northeast of Sunyani (Regional Capital). It lies within latitudes 7° 30' and 8° 05' North and longitudes 2° 15' West and 1° 55' East. In terms of land area, Wenchi Municipal covers 3,494 sq kilometers. The Municipality shares boundaries with Techiman Municipal to the west, Kintampo Municipal on the northwest, Tain district to the east and Sunyani Municipal to the south. Wenchi, the Municipal capital is 29km from Techiman. Most of the indigenes in the municipality are peasant farmers, whereas only a few are into large scale farming. The major crops grown in the municipality include Root and Tuber (yam, cassava and cocoyam), cereals and legumes (maize, groundnut, cowpea, soya bean, bambara and sorghum), Vegetables (okro,

pepper, garden eggs, tomato, water melon), Cash crops (cashew, mango, cocoa, citrus and oil palm), and others such as plantain.

### **1.6 Data source and organization**

The data used for the study is purely secondary, obtained from the Education Management and Information System (EMIS) department of the Wenchi municipal directorate of education. Microsoft Excel software was used for data entry as well as data organization. It was also used to edit the data for consistency.

### **1.7 Data analysis**

The statistical technique which was used to analyze the data is the Generalized Estimating Equations (GEE) family of models. The statistical softwares to be employed for analyzing the data are R and SPSS.

#### *Justification for using GEE:*

The response variable, enrolment, are counts which would be regressed on other several covariates some of which are also counts, binary and even categorical. Thus Poisson regression models (GLM's) are appropriate, but due to the repeated measurements on the schools, there is a higher likelihood of association among the repeated measurements (i.e. within schools). Due to this the Poisson regression models becomes inappropriate. Also, coupled with the fact that the study seeks to find out the overall effect due to changes in the predictors on the whole population, the GEE family of models would be most appropriate.

### **1.8 Significance of the Study**

Although there are several studies that have examined enrolment trends in basic schools as a result of the effects of capitation grant and school feeding, there are only a few or no studies that have researched into the effects on enrolment using Generalized Estimating Equations (GEE).

The findings from this research work will go a long way in aiding educationists and educational policy makers to become fully aware of the variables that have direct influence on enrolments in basic schools.

To a very large extent, it would contribute to knowledge since little has been done using this technique with educational surveys.

### **1.9 Organization of the Study**

The rest of the thesis is organized into four chapters, chapter two comprised of reviews of related literature on the topic and chapter three looks at the GEE family of models under different working correlation assumptions and methods of estimating the parameters. Chapter four analyzes and discusses the findings of the study and finally, chapter five presents the discussions, conclusions and recommendations of the study.

## **CHAPTER TWO**

### **LITERATURE REVIEW**

#### **2.0 Introduction**

This chapter discusses the literature available on the Capitation grant and the School Feeding Programme, as well as their impact on key educational indicators.

#### **2.1 Feeding in Basic Schools**

Ghana, like many other developing countries in Africa faces socio-economic problems such as high poverty rates, high illiteracy rates, and low standards of education, among others. It is an undeniable fact that standards of education gradually falls when basic education which is supposed to be the foundation is challenged with enormous problems. Few among these problems are poor feeding and nutrition, health of school children and high dropout rate. Several research works have shown that feeding and proper nutrition have significantly imparted on some key education outcomes and indicators. Advocates of child health have experimented with students' diets in the United States for more than twenty years. Initial studies focused on benefits of improving the health of students are apparent. Likewise, improved nutrition has the potential to positively influence students' academic performance and behavior. Though researchers are still working to definitively prove the link, existing data suggests that with better nutrition students have fewer absences, and students' behavior improves, causing fewer disruptions in the classroom (Sorhaindo and Feinstein, 2006).

### **2.1.1 Brain functioning in children**

Several studies show that nutritional status can directly affect mental capacity among school-aged children. For example, iron deficiency, even in early stages, can decrease dopamine transmission, thus negatively impacting cognition (Pollitt, 1993). Deficiencies in other vitamins and minerals, specifically thiamine, vitamin E, vitamin B, iodine, and zinc, are shown to inhibit cognitive abilities and mental concentration (Chenoweth, 2007). Additionally, amino acid and carbohydrate supplementation can improve perception, intuition, and reasoning (Frisvold, 2012). There are also a number of studies showing that improvements in nutrient intake can influence the cognitive ability and intelligence levels of school-aged children (Eysenck & Schoenthaler, 1997).

### **2.1.2 Basic school Enrolment, attendance and Retention**

The word “enroll” means to put yourself or someone else onto the official list of members of a course, college or group. Attendance is the ability of the student to show up in school for every school going day whereas primary retention rate is a ratio of the total number of children who are able to complete primary six as against the total number of children enrolled in primary one, expressed as a percentage. Previous works have shown that good nutrition helps students show up at school prepared to learn. Because improvements in nutrition make students healthier, students are likely to have fewer absences and attend class more frequently. Thus, student attendance improves in schools that implement universal free school breakfast programs (Wahlstrom and Begalle, 1999).

## **2.2 Ghana's School Feeding Programme (GSFP)**

On the basis of these enormous research works on how nutrition and feeding in basic schools positively affects educational outcomes, the Government of Ghana, with support from the Dutch Government commenced the implementation of the GSFP in 2005. The GSFP is an initiative under the Comprehensive African Agricultural Development pillar III which seeks to enhance food security and reduce hunger. It is also in line with the UN MDG's since it seeks to reduce hunger and poverty.

The GSFP's formal objective is to work on a long term solution for poverty in Ghana. It addresses food security by aiming to provide one hot nutritious meal a day for children in public primary and kindergartens in the poorest areas of the country, using locally grown foodstuffs. However, the medium term objectives of the GSFP includes; to increase school enrolment, attendance and retention; reduce hunger and malnutrition; and to boost domestic food production.

Due to inadequate funding, the feeding could not extend to all basic schools throughout the country, and so commenced with only 10 pilot schools selected from each region of the country. By August 2006, the number of schools had been increased to 200 covering about 69,000 pupils in 138 districts, (ECASARD/SNV Ghana, May 2009).

### **2.3 School Fees Abolition and the Capitation Grant**

Countries worldwide are making conscious and remarkable efforts towards reducing the number of out-of-School children. School fees often prevent children in developing countries from being able to go to school. It has been one of the biggest barriers in the expansion of schooling in the poorest countries. Experiences in several countries show that the private cost of schooling to households prevents children from accessing and completing quality basic education. No child should be excluded from schooling because of a family's inability to pay, but poverty often imposes tough choices on families and households about how many children to send to school, which children to send to school and how long they may attend. For instance, a survey carried out in 2004 found that households in Zambia finance 50% to 75% of total primary education spending and in Malawi; the average household expenditure for public primary school was nearly 80% (Ampratwum and Armah-Attoh, 2004). Moreover, the strides made over the last decade to get millions of children into school are nothing short of outstanding. The number of children who are out of school has decreased significantly – from 115 million in 2001 to 67 million in 2009. However, global progress towards universal education has slowed since 2005 and based on current trends, the out-of-school population could increase to 72 million by 2015, which means the Millennium Development Goal of universal primary education for all will not be achieved.

These developments prompted world leaders to explore ways and means of improving their education system in order to achieve their commitment to education for all. But in recognition of poverty as the major stumbling block to achieving universal primary education, the School Fee Abolition Initiative (SFAI) was launched in 2005 by UNICEF and the World Bank as one of the

‘Bold Initiatives’ aiming to make a breakthrough in access to basic education and significantly scaling up progress to meet the MDGs and EFA targets in the next decade. Countries that have taken the bold step to eliminate fees saw a dramatic and sudden surge in enrolment as a result: In Uganda in 1996, primary school enrolment grew from 3.4 million to 5.7 million; and in Kenya in 2003, enrolment increased from 5.9 million to 7.2 million.

Ever since the implementation of the free schooling initiative, many have expressed that it has been the most important policy measure that has had a dramatic, transforming impact on school enrolment so far, as it unleashes latent demand for education and encourages children from disadvantaged backgrounds to participate. One of such comments made by a commonwealth official is quoted as *“We realize the role that the School Fee Abolition Initiative can play in boosting enrolment among the poor. It has a special relevance for girls, as they are more adversely affected by poverty. It is especially significant for us, as the Commonwealth is estimated to be housing about two thirds of the world's out-of-school children”*, (Jyotsna Jha, Commonwealth Secretariat).

Ghana is no exception as far as the school fees abolition initiative is concerned. As part of strategies to achieve universal primary education, the Government of Ghana launched the Capitation grants policy in 2005 where every public basic schools including kindergarten, primary and junior high schools were paid GH¢3.00 (Three Ghana cedis) for every child enrolled in school. The capitation grant, meant to replace school fees, is to help equip schools with the relevant resources/materials needed to carry out school quality improvement activities and mainly to boost enrollment drive in basic schools. Moving forward, the grant was later

increased to GH¢4.50 (Four Ghana cedis, and fifty peswas) in 2009 to account for inflationary effects (Ampratwum and Armah-Attoh, 2004).

## **2.4 Evaluation of Educational Interventions**

After the implementation of the School Feeding Programme and the Capitation grant policy in Ghana, there has been several studies conducted to assess and evaluate the real impacts these policies have had on key educational outcomes such as enrolment trend, attendance, retention, gender equity and performance. Other studies have also been conducted to highlight the factors hindering the smooth implementation of these policies (the problems) and their possible remedy.

Osei et al. (2009) in their work “Effects of capitation grant on education outcomes in Ghana” employed econometric estimation model to assess how capitation grant affects gross enrolment rates at the Junior High school level, the pass rates for the national examinations at the Junior High school level and the gap in the examination performance of boys and girls, using 138 district level data over the period 2005- 2007 and across the country. The results of the study revealed that; the capitation grant has not had significant impact on BECE pass rates in Ghana, no significant relationship existed between capitation grant and gross enrollment, and capitation grant has not impacted on bridging the gap between the BECE pass rates for male and females. The researcher however noted that although the results are not consistent with a priori expectations, the findings may reflect the fact that the capitation grant in Ghana only started in 2005 and so it is too early to begin to see its effects.

Adding to that, the Ministry of Education in their annual report in 2010 indicates that the introduction of poverty alleviation interventions such as the capitation grant, the school feeding programme, free exercise books and uniforms have increased enrolment at all levels. The report further indicates that the major challenges in the achievement of universal primary education are the out of school children and those dropping out of schools. Moreover, education quality declined as the provision of teachers, facilities and logistics lagged behind the increase in enrolment. Thus, giving the statistics for the period; GER was quoted as 94.9%, Primary Completion Rate as 87.1% and BECE Pass Rate as 62.42% (Education Sector Performance Report, 2010).

Other research works were done in 2011 in relation to the subject under discussion. Firstly, Osei-Fosu (2011) in his work “Evaluating the impact of the school feeding programme and the capitation grant on enrolment, attendance and retention in schools” (The case of Weweso circuit) used difference-in-difference method by comparing changes in enrollment, attendance and retention between before and after, and between beneficiary schools (treatment) and non-beneficiary schools (control). The researcher also employed simple regression analysis to find the impact of these two policies on enrollment, attendance and retention for the 2001/2002 and 2008/2009 academic years. Data for the study were obtained by randomly selecting 20 basic schools from the Weweso circuit of the Asokwa Sub-Metro Education district in the Kumasi Metropolis, and the study concludes that the capitation grant has had an impact on enrollment (though not significant) and has also had significant impact on attendance and retention. It also came out that the school feeding programme had high positive and significant impact on school enrollment, attendance and retention.

Again, within the same year, a survey by Asante Comfort and titled “The Capitation Grant: Impact on enrollment of pupils in the Basic Education Schools in Ghana” was done by randomly selecting 100 respondents from 20 basic schools within the Sunyani Metropolis. It used a survey to assess the opinions of head teachers, teachers and other stakeholders of education regarding the impact of the capitation grant policy on pupils’ enrollment in the municipality. The findings of the study revealed that the capitation grant has actually led to increase in the enrollment of pupils in the basic schools.

Finally, Dawuda (2012), in his work “The Impact of Capitation Grants on Access to Primary Education in Ghana”, determined the extent to which the capitation grants policy has contributed to increase access and participation in primary education. Data for the study was obtained from the World Development Indicators, UIS and Ghana Education Management Information System (EMIS). The analysis compared the key educational indicators such as the Gross Admission Rate (GAR), Net Admission Rate (NAR), out-of-primary-school children, Gross Enrolment Rate (GER), Net Enrolment Rate (NER), Gender Parity Index (GPI) and primary completion rate, five years before implementation of capitation grants and five years afterwards. The results of the study showed significant improvement in enrollment following the implementation of the capitation grant policy in Ghana; although the increases were inadequate to enable Ghana attain the education-related MDG by the target date by 2015.

## **2.5 Modeling Longitudinal Data**

Longitudinal studies are increasingly common in many scientific research areas. The longitudinal data (sometimes referred to as panel data) are defined as the data resulting from the observations

of subjects (human beings, animals, or laboratory samples, etc.) which are measured repeatedly over time (Bijleveld et al., 1999). The purpose of conducting longitudinal study is to look at; how treatment means differ, how treatment means change over time and how differences between means of treatments change over time period. When change itself is the object of study, the only way to investigate the change is by collecting repeated measurement (Michikazu and Weiming, 2009). It is worth noting that since longitudinal data involves the measurement of the same subject repeatedly (repeated measures), the observations are not independent.

There are a number of statistical models that can be used for longitudinal data analyses. First is the univariate and multivariate analysis of variance (ANOVA and MANOVA for short). Both models assume interval measurement and that the errors are normally distributed and homogeneous across groups. The weak aspect of these methods is that they only estimate and compare the group means and not informative about individual growth. Furthermore, as an assumption, these methods must have fixed time points (Michikazu and Weiming, 2009).

Again, another statistical technique suitable for longitudinal studies is the Mixed-effect Regression model (MRM), which clearly models individual change across time. The MRM has an added advantage of handling subjects measured at different time points and hence can also be used for incomplete longitudinal data. It does not also require restrictive assumptions concerning missing data across time.

## **2.6 Marginal Modeling**

For longitudinal data, measurements of the same subject are taken repeatedly over time, so therefore observations are not independent, but rather correlated. As such these correlations

needed to be taken into account in modeling; otherwise the standard errors of the estimates would be underestimated for the between-subject and overestimated for the within-subject effects.

Again, because the independence assumption is violated in longitudinal data analyses, it mostly uses quasi-likelihood estimation, and so the full likelihood of the data is not specified.

Marginal models tend to model the regression of the response on a set of explanatory variables, and the within-subject dependence (i.e., the association parameters) separately. In GEE models, the term “marginal” indicates that the model for the mean response depends only on the covariates of interest, and not on any random effects or previous responses (Fitzmaurice et al., 2004).

## **2.7 Generalized Estimating Equations (GEE)**

Generalized Linear Models (GLMs) represent a class of models that are used to fit fixed effects regression models to normal and non-normal data (McCullagh & Nelder, 1989). GEEs, introduced by Liang and Zeger (1986) are extensions of the GLM to longitudinal (correlated) data using quasi-likelihood estimation. There is no convenient or natural specification of the joint multivariate distribution of when the responses are discrete, and thus GEE is a popular alternative to Maximum Likelihood (ML) estimation which provides a general approach for analyzing discrete and continuous responses with marginal models.

A basic premise of GEE approach is that one is primarily interested in the regression parameter and is not interested in the variance-covariance matrix of the repeated measures. As such, generalized estimating equations treat the covariance structure as a nuisance and they are not

concerned about variance of each data (Michikazu and Weiming, 2009). By relying on the independence across subjects, GEE method yields consistent estimates of regression coefficients and their variances (thus, standard errors), even when the structure of the covariance matrix is mis-specified or incorrect.

Owusu-Darko (2011) applied GEE model on student's academic performance using all of the 126 students in the mathematics department of the KNUST. Questionnaires were used, in addition to the existing data of students Semester Weighted Average (SWA) records, to solicit their opinion about factors that affect their performance in mathematics. The variables used were gender, entry age into the school, geographical location and graded level of former school attended. At the end of the study, the coefficient estimation of the study parameters revealed that only geographical location of students is significant and therefore affects their academic performance.

## CHAPTER THREE

### METHODOLOGY

#### 3.0 Introduction

This chapter explains into details the statistical technique used to analyze the data. The data for this research work is longitudinal in nature.

Longitudinal data is a type of repeated measures where outcomes are measured repeatedly overtime, such that its analysis allows for evaluation of change overtime. In repeated measures designs, there are several individuals and measurements are taken repeatedly an each individual. When these repeated measurements are taken over time it is called a longitudinal study or, in some applications, a panel study.

Sometimes researchers also use longitudinal data interchangeably with cross-sectional data but, there is a clear distinction between the two. In many branches of science, studies are often designed to investigate changes in a specific parameter which is measured repeatedly overtime in the participating subject. Data for such studies are called longitudinal data, in contrast to cross-sectional data where the response of interest is measured only once for each individual.

The major advantage of longitudinal studies over cross-sectional studies is that it helps to know the change overtime because information is collected at different time points; so you can adjust for the variability between individuals. Diggle et al. (2002) pointed out that one of the main advantages of longitudinal studies is its ability to distinguish changes over time within individuals from differences among people in their baseline values.

Again, the number of subjects needed for analysis increase with cross-sectional study. For instance, if  $n$  subjects are needed in longitudinal studies, then  $3n$  subjects will be needed for cross-sectional studies.

However, the major shortcoming of longitudinal studies includes a longer follow time period for each individual. As a result of this longer follow time period, there is the potential problem of missing data. A substantial number of missing data makes longitudinal studies very complicated and complex. The design for longitudinal studies are in two folds; the prospective and the retrospective designs.

With prospective randomized (observational studies), designs subjects are followed through to the end of the studies, and the researcher is at least assured of a better data quality because at the design stage he/she decides, with careful considerations, which time points to select or to measure the response in order not to end up with a lot of missing data. Additionally there is sometimes the possibility of follow-up measures through for example, phone calls.

But with retrospective designs on the other hand, the researcher often does not have much control because a lot of times the data is already collected in a data base and one is compelled to use what is available in the data base. A major problem with this design is that there is a high risk of missing data and in most case becomes virtually impossible to make follow-ups on subjects to obtain the additional information.

The choice of models for analyzing longitudinal data depends on the scientific question of interest. That is, appropriate analysis for longitudinal data depends on the study objectives and

design. It is however worth noting that the analysis of longitudinal data is very complex especially in the presence of missing data, categorical responses and time – varying covariates.

However, the choice of appropriate model should be one that accounts for correlations between measurements from the same subject; because ignoring their correlations may lead to incorrect inferences, bias results or less precise estimates. Again, the appropriateness of the chosen model should be able to account for missing data, since its failure may lead to incorrect conclusion.

Many of the earliest models for the analysis of change were based on the Analysis of variance (ANOVA) originally developed by Fisher. The mixed-effect ANOVA model (also called the univariate repeated- measures ANOVA) with a single random subject effect, was one of the earliest models used for analyzing longitudinal data. The main rationale for the inclusion of a random subject effect is to induce positive correlation among the repeated measurements taken on the same subject. Statisticians recognized that a longitudinal data structure with  $N$  individuals and  $n$  repeated measurements has striking similarities to data collected in a randomized block design, and therefore seemed appropriate to apply ANOVA methods to the repeated – measures data collected from longitudinal studies. In this way, the  $N$  individuals are regarded as the blocks/main plots whereas the  $n$  repeated measurements are taken as treatments.

The univariate repeated measures ANOVA model can be written as;

$$Y_{ij} = X'_{ij}\beta + b_i + e_{ij}, \quad i=1..N, \quad j=1..n$$

where

$Y_{ij}$  - outcome of interest (response)

$\mathbf{X}_j$  - design vector (vector of covariates)

$\boldsymbol{\beta}$  - vector of fixed effect regression parameters

$b_j$  - random effects

$e_j$  - error terms, and also

$$b_j \sim N(0, \sigma_b^2), e_j \sim N(0, \sigma_e^2)$$

The random effect  $b_j$  represents an aggregation of all the unobserved or unmeasured factors that make individuals respond differently.

As situation advanced, a related approach for the analysis of longitudinal data, with an equally long history, but requiring a more advanced computations was introduced. This model is just the multivariate analogue of the univariate repeated-measures ANOVA case, known as the multivariate repeated-measures ANOVA. The only difference here is that the MANOVA is a model for multivariate responses. Statisticians recognized the multivariate repeated measures ANOVA model for the analysis of longitudinal studies because its data structure exhibits a multivariate vector of substantively distinct response variables measured repeatedly on the same subjects over a period of time.

However, the two ANOVA – based approaches described as models for analyzing longitudinal data has shortcomings that limited their usefulness in their applications. The univariate repeated measures ANOVA made very restrictive assumptions about the covariance structure for repeated measures on the same individual. For example, the constraint on the correlation among repeated measurements is somewhat unappealing for longitudinal data, where the correlations are

expected to decay with increasing separation in time. Moreover, the assumption of constant variance across time is often unrealistic. In many longitudinal studies the variability of the response at the beginning of the study is discernibly different from the variability towards the completion of the study.

In contrast, the multivariate repeated measures ANOVA did not make restrictive assumptions on the covariance among the longitudinal responses on the some individual. This implies that correlations could assume any pattern where as the variability could also change overtime. However, the repeated measures MANOVA cannot be used when the design is unbalanced over time that is when the vector of repeated measures are of different lengths and/or obtained at different time points. Another problem of this model is its inability to allow for general missing data patterns to arise. The severity of this problem is that the entire data vector for an individual is excluded from the analysis even if the individual has a single missing response at any occasion the can produce biased estimators of change in the mean response overtime.

It is clearly evident that data which mostly results from longitudinal studies are highly unbalanced and not readily amenable to ANOVA methods developed for balanced designs. This led to the development of more robust and versatile models that can handle the commonly encountered problems of data that are unbalanced and incomplete, mistimed measurements, time varying and time invariant covariates and responses that are discrete rather than continuous. Some of these flexible models developed thereafter are conditional and marginal models. In order to describe these flexible models for correlated data, it is important to understand a few fundamental concepts. These are treated in the next few sections.

### 3.1 The Exponential Family of Distributions

#### 3.1.1 Introduction

The exponential family of distributions is one of the most important classes of distributions in statistics. It represents a unified natural set of probability distributions of special forms, chosen for mathematical convenience, on an account of some useful algebraic properties as well as for generality. They include many of the most common distributions such as the normal, exponential, gamma, beta, chi squared, Bernoulli, Poisson and many others. The concept of exponential families is credited to Pitman et al., (1936).

*Definition:*

Let  $Y$  be a single random variable whose probability distribution depends on a single parameter  $\theta$ , then the distribution of  $Y$  belongs to the exponential family if it can be written in the form;

$$f(y:\theta) = s(y) \exp\{t(\theta) a(y) b(\theta)\} \quad (3.1)$$

where  $a$ ,  $b$ ,  $s$  and  $t$  are known functions.

Equation (3.1) can further be rewritten in the form;

$$f(y:\theta) = \exp\{a(y)b(\theta) + c(\theta) + d(y)\} \quad (3.2)$$

where

$$d(y) = \ln\{s(y)\} \text{ and } c(\theta) = \ln t(\theta)$$

Also, if  $a(y) = y$ , the distribution is said to be in the canonical (standard) form.

$b(\theta)$  is the natural parameter of the distribution. Actually, the natural parameter,  $b(\theta)$  becomes the link function in GLM's

However, if there are other parameters in addition to the parameter  $\theta$ , they are regarded nuisance parameters and are treated as if they are known. As noted early on, distributions such as the Normal, Binomial/Bernoulli and Poisson are all members of the exponential family. Members of this family often share some common properties.

### 3.1.2 Properties of the exponential family

Here expression for the expected value and variance of  $a(y)$  would be found, provided that the order of integration and differentiation can be interchanged on  $f(y; \theta)$ .

It is also known that if  $f(y; \theta)$  is a probability density function (pdf), then;

$$\int_{dy} f(y; \theta) dy = 1 \quad (3.3)$$

Differentiating equation (3.3) with respect to  $\theta$  gives

$$\frac{d}{d\theta} \int f(y; \theta) dy = 0$$

(3.4)

Reversing the order of integration and differentiation, equation (4) becomes

$$\int \frac{d}{d\theta} f(y; \theta) dy = 0$$

(3.5)

Similarly, the second order differential with respect to  $\theta$  and reversing the order of integration in equation (3.3) gives

$$\int \frac{d^2}{d\theta^2} f(y; \theta) dy = 0 \quad (3.6)$$

Again, we know that

$$f(y:\theta) = \exp\{a(y)b(\theta) + c(\theta) + d(y)\}$$

so

$$\frac{df(y:\theta)}{d\theta} dy = [a(y)b'(\theta) + c'(\theta)] \exp\{a(y)b(\theta) + c(\theta) + d(y)\}$$

which is simplified as

$$\frac{df(y:\theta)}{d\theta} dy = [a(y)b'(\theta) + c'(\theta)] f(y:\theta) \quad (3.7)$$

Then from equation (3.5)

$$\int [a(y)b'(\theta) + c'(\theta)] f(y:\theta) dy = 0$$

$$b'(\theta) \int a(y) f(y:\theta) dy + c'(\theta) \int f(y:\theta) dy = 0$$

From definition of expectation,

$$\int a(y) f(y:\theta) dy = E[a(y)]$$

so we shall have

$$b'(\theta) E[a(y)] + c'(\theta) = 0$$

hence we have the expectation as

$$\mu = E[a(y)] = \frac{-c'(\theta)}{b'(\theta)} \quad (3.8)$$

A similar argument can be used to obtain  $\text{Var}[a(y)]$ . Differentiating equation (3.7) again with

respect to  $\theta$  gives

$$d^2 f(y:\theta) dy = [a(y)b''(\theta) + c'(\theta)] f(y:\theta) + [a(y)b'(\theta) + c'(\theta)]^2 f(y:\theta)$$

From equation (3.6)

$$\int [a(y)b''(\theta) + c'(\theta)] f(y:\theta) dy + [a(y)b'(\theta) + c'(\theta)]^2 f(y:\theta) dy = 0$$

which can be simplified as

$$b' \int a(y) f(y:\theta) dy + c'(\theta) \int f(y:\theta) dy + (b'(\theta))^2 \int a(y) + \frac{c'(\theta)}{b'(\theta)} dy = 0$$

so from equation (3.8) we have

$$b'E[a(y)] + c'(\theta) + [b'(\theta)]^2 \int a(y) - E[a(y)] dy = 0$$

and also by definition,

$$\int a(y) - E[a(y)] dy = \text{Var}[a(y)]$$

by substituting, we have

$$b'E[a(y)] + c'(\theta) + [b'(\theta)]^2 \text{var}[a(y)] = 0$$

simplifying, we finally have

$$\text{var}[a(y)] = \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{[b'(\theta)]^3} \quad (3.9)$$

There are a group of statistical models which stems from the exponential family, and shares its desirable properties as well. These models are extensions of the traditional linear models for continuous and non-continuous outcomes.

## 3.2 Generalized Linear Models (GLM's)

### 3.2.1 Introduction

The Generalized linear model, used for outcomes that are members of the exponential family, was first introduced by Nelder and Weddeburn (1972). They provided a unified framework to study various regression models, rather than a separate study for each individual regression. GLM's are extensions of the classical linear model for situations of independent continuous, counts, binary and categorical outcomes. They include linear regression models, analysis of variance models, logistic regression models, Poisson regression models, log linear models, among others.

*Definition:*

Let  $Y_1, Y_2, \dots, Y_N$  be a set of independent random variable each with a distribution from the exponential family such that  $E(Y_i) = \mu_i$  where  $\mu_i$  is some function of  $\theta_i$ , then,

- (i) The distribution of each  $Y_i$  has the canonical form and depends on a single parameter  $\theta$ , thus;

$$f(y_i; \theta) = \exp\{y_i b(\theta) + c_i(\theta) + d_i(y_i)\}$$

- (ii) The distribution of all the  $Y_i$ 's are of the same form (example Binomial, Normal, Poisson) so that the subscripts on  $b$ ,  $c$  and  $d$  are not important. In this case, the joint probability density function which is

$$f(y_1, y_2, \dots, y_N; \theta_1, \theta_2, \dots, \theta_N) = \prod_{i=1}^N \exp\{y_i b(\theta_i) + c_i(\theta_i) + d(y_i)\}$$

then becomes

$$\exp\left[\sum_{i=1}^N y_i b(\theta) + \sum_{i=1}^N c(\theta) + \sum_{i=1}^N d(y_i)\right]$$

If the  $Y_i$ 's satisfy the above properties, then the generalized linear model has three major components;

- (i) Random component: A response variable  $Y_1, Y_2, \dots, Y_N$  which are assumed to share the same distribution from the exponential family. This therefore specifies the distribution of the response variable.
- (ii) Systematic component: A linear combination of the predictor variables and the regression coefficients in the form

$$\eta = \beta_0 + \beta_1 X_1 + \dots + \beta_j X_j = \mathbf{X}\beta$$

The explanatory variables may be quantitative (continuous), qualitative (discrete) or both.

- (iii) Link function: A smooth and invertible linearizing link function  $g(\cdot)$  which transforms the expectation of the response variable  $\mu = E(Y_i)$  to the linear predictors. It can be written as

$$g(\mu) = \eta = \mathbf{X}\beta \text{ where } \mu = E(Y_i)$$

because the link function is invertible, we can also write

$$\mu = g^{-1}(\eta) = g^{-1}(\mathbf{x}\beta)$$

### 3.2.2 GLM for Count Data

#### *Introduction*

A great deal of data collected by scientists, medical statisticians and economists, is in the form of counts. Typically, these are the numbers of occurrences of some event in a defined time period or space. Examples include the number of medical conditions reported by a person, the number of spelling mistakes on the page of a newspaper and the number of faulty components in a batch of manufactured items. Data in the form of counts are usually modeled using the Poisson distribution.

The Poisson probability distribution was first introduced by Simeon-Denis Poisson in his work “Research on the probability of judgments in criminal and civil matters” (1838). The distribution is often used to model count data, by focusing on a random variable,  $Y$  that counts, among other things the number of discrete occurrences that take place during a time interval. If the expected (average) number of occurrences is  $\theta$ , then the probability that there are exactly  $y$  occurrences is given by

$$f(y:\theta) = \frac{\theta^y e^{-\theta}}{y!}, \quad y=0, 1, 2, \dots \quad (3.10)$$

However, before the Poisson distribution can be used, certain assumptions must be met to ensure its suitability. The major assumptions underlying the Poisson distribution are enumerated below:

- (i) Observations are independent
- (ii) Probability of another occurrence in such a short interval is zero.

(iii) Probability of occurrence in a short interval is proportional to the length of the interval.

*Model specification*

The Poisson regression model is a technique used to describe count data as a function of a set of predictor variables. Let  $Y_i$  be a vector of count response random variable, where  $i=1,2,\dots,n$ ,  $\mathbf{X}$  be a matrix of predictors and  $\boldsymbol{\beta}$  be a vector of unknown parameter coefficients.

The Pdf in equation (3.10) can be re-written in the exponential form as

$$f(y:\theta) = \exp\{y \ln \theta - \theta - \ln y!\}$$

comparing with equation (2),

$$a(y) = y, \quad b(\theta) = \ln \theta, \quad c(\theta) = -\theta \text{ and } d(y) = \ln y!$$

Again, since the Poisson distribution is a member of the exponential family, its mean and variance can be derived from the properties of this family of distributions.

Now from above,

$$b(\theta) = \frac{1}{\theta}, \quad b'(\theta) = -\frac{1}{\theta^2}, \quad c'(\theta) = -1, \quad c''(\theta) = 0$$

From equation (8)

$$E(y) = \frac{-(-1)}{-1/\theta}$$

$$E(Y) = \theta$$

Also from equation (9)

$$\text{var}(Y) = \frac{\left(-\frac{1}{\theta}\right)(-1) - 0}{\left(\frac{1}{\theta}\right)}$$

$$\text{var}(Y) = \left(\frac{1}{\theta}\right) \left(\frac{\theta}{1}\right)$$

$$\text{Var}(Y) = \theta$$

Hence for Poisson distribution,  $E(Y) = \text{var}(Y) = \theta$

Since  $a(y) = y$ , then  $Y$  has the canonical form, and also  $b(\theta) = \ln \theta$ , has the log link as the natural parameter. The expected value of the mean response is linked to the systematic component through the log link since Poisson regression is a GLM.

That is

$$E(Y) = \mu = g^{-1}(\eta)$$

$$\ln \theta = X\beta \tag{3.11}$$

which then gives

$$\hat{\theta} = e^{X\beta} \tag{3.12}$$

The above equation is called the Poisson regression model.

### 3.3 Estimation method for GLM's

Parameter estimation in GLM's is carried out mainly by maximum -likelihood estimation and in some cases by quasi -likelihood estimation.

#### 3.3.1 Maximum- likelihood Estimation (GLMs)

Consider independent random variables  $Y_1, \dots, Y_N$  satisfying the properties of a generalized linear model. We wish to estimate parameters  $\beta$  which are related to the  $Y_i$ 's through

$$E(Y_i) = \mu \text{ and } g(\mu) = X_i \beta$$

where  $\mathbf{X}_i$  is a vector of covariates. For each  $Y_i$ , the log-likelihood function is

$$\ell_i = y_i b(\theta) + c(\theta) + d(y_i) \quad (3.13)$$

Also from

$$\mu = E(Y_i) = -c'(\theta) / b'(\theta),$$

$$\text{var}(Y_i) = \frac{b''(\theta)c'(\theta) - b'(\theta)c''(\theta)}{[b'(\theta)]^3}$$

and

$$g(\mu) = x_i' \beta = \eta \quad (3.14)$$

The log-likelihood function for all the  $Y_i$ 's is

$$\ell = \sum_{i=1}^n \ell_i = \sum y_i b(\theta) + \sum c(\theta) + \sum d(y_i)$$

To obtain the maximum likelihood estimator for the parameter  $\beta_j$  we need

$$\frac{\partial \ell}{\partial \beta_j} = \mu_j = \sum_{i=1}^n \left( \frac{\partial \ell_i}{\partial \beta_j} \right) = \sum_{i=1}^n \left( \frac{\partial \ell_i}{\partial \theta} \cdot \frac{\partial \theta}{\partial \mu} \cdot \frac{\partial \mu}{\partial \beta_j} \right) \quad (3.15)$$

That is, using the chain rule for differentiation. From equation (3.13),

$$\frac{\partial \ell_i}{\partial \theta} = y_i b'(\theta) + c'(\theta)$$

Also from equation (3.8)

$$-b'(\theta) E(y_i) = c'(\theta)$$

so

$$\frac{\partial \ell_i}{\partial \theta} = y_i b'(\theta) - b'(\theta) E(Y_i)$$

which simplifies to

$$\begin{aligned}\frac{\partial \ell_i}{\partial \theta} &= b(\theta) [y_i - E(Y_i)] \\ &= b(\theta) [y_i - \mu]\end{aligned}\tag{3.16}$$

$$\text{Also } \frac{\partial \theta}{\partial \mu} = \frac{1}{\left(\frac{\partial \mu}{\partial \theta}\right)}$$

So from equation (3.8) again,

$$\begin{aligned}\frac{\partial \mu}{\partial \theta} &= \left[ \frac{b(\theta)c'(\theta) - c'(\theta)b'(\theta)}{[b(\theta)]^2} \right] \\ &= \frac{b'(\theta)c'(\theta) - b(\theta)c''(\theta)}{[b(\theta)]^2}\end{aligned}\tag{3.17}$$

Multiply equation (3.17) by  $\frac{b(\theta)}{b(\theta)}$ , we have

$$\frac{\partial \mu}{\partial \theta} = \frac{b(\theta) [b'(\theta)c'(\theta) - b(\theta)c''(\theta)]}{[b(\theta)]^3}$$

Hence from equation (3.9)

$$\frac{\partial \mu}{\partial \theta} = b'(\theta) \text{var}(y_i)\tag{3.18}$$

Finally from equation (3.14)

$$\frac{\partial \mu}{\partial \beta_j} = \frac{\partial \mu}{\partial \eta} \cdot \frac{\partial \eta}{\partial \beta_j}, \quad \frac{\partial \eta}{\partial \beta_j} = x_{ij}$$

$$\frac{\partial \mu}{\partial \beta_j} = \left( \frac{\partial \mu}{\partial \eta} \right) x_{ij} \quad (3.19)$$

Hence substituting equations (3.16), (3.18) and (3.19) into equation (3.15), the score equation for GLM is given by

$$T_j = \sum_{i=1}^n b'(\theta)(y_i - \mu) \cdot \frac{1}{b(\theta) \text{var}(y_i)} \cdot x_{ij} \left( \frac{\partial \mu}{\partial \eta} \right) = 0$$

which finally simplifies to

$$T_j = \sum_{i=1}^n \left[ \frac{(y_i - \mu)}{\text{var}(Y_i)} x_{ij} \left( \frac{\partial \mu}{\partial \eta} \right) \right] = 0 \quad (3.20)$$

### 3.3.2 Quasi – Likelihood Estimation

In many problems of statistical estimation, we know some detail of the distribution governing the data, but may be unwilling to specify it exactly. This limits the usage of maximum likelihood which requires exact /full specification of the distribution so as to construct the likelihood and subsequently the log-likelihood. The idea of quasi – likelihood developed by Wedderburn (1974), addresses this concern.

Thus, quasi – likelihood method of estimation requires only a model for the mean of the data, and the relationship between the mean and the variance. In many cases, these methods of estimation retain the full or nearly full efficiency compared to the maximum likelihood.

The quasi – likelihood  $Q(y_i; \mu)$  is defined as

$$\frac{\partial Q(y_i; \mu)}{\partial \mu} = \frac{y_i - \mu}{V(\mu)}$$

### 3.3.3 Parameter estimation for Poisson regression model

The Poisson regression model uses the method of maximum likelihood to estimate the model coefficients ( $\beta$ ). From the probability density function in equation (10), the likelihood function is given as

$$L = \prod_{i=1}^n f(y; \theta) = \prod_{i=1}^n \frac{\theta^y e^{-\theta}}{y!}$$

Then the log-likelihood is given as

$$\ell = \ln \prod_{i=1}^n f(y; \theta) = \ln \prod_{i=1}^n \frac{\theta^y e^{-\theta}}{y!}$$

Since logarithms converts products into sums, we have

$$\ell = \sum_{i=1}^n \ln \left( \frac{\theta^y e^{-\theta}}{y!} \right)$$

$$\ell = \sum_{i=1}^n (y_i \ln \theta - \theta - \ln y_i!)$$

$$\ell = \sum_{i=1}^n y_i \ln \theta - \sum_{i=1}^n \theta - \sum_{i=1}^n \ln y_i!$$

$$\ell = \sum_{i=1}^n y_i x_i \beta - \sum_{i=1}^n e^{x_i \beta} - \sum_{i=1}^n \ln y_i!$$

(3.21)

Differentiating the log likelihood function in equation (3.21) with respect to  $\beta$ , we have

$$\frac{\partial \ell}{\partial \beta} = \sum_{i=1}^n y_i x_i' - \sum_{i=1}^n x_i' e^{x_i \beta}$$

$$= \sum_{i=1}^n (y_i - e^{x_i \beta}) x_i'$$

To obtain maximum values of  $\beta$ , we set

$$\sum_{i=1}^n (y_i - e^{x_i \beta}) x_i = 0 \quad (3.22)$$

Equation (3.22) is referred to as the score equation.

The second differential of the log likelihood function in equation (3.21) gives the Hessian, given by

$$\frac{\partial^2 \ell}{\partial \beta \partial \beta_j} = - \sum_{i=1}^n (e^{x_i \beta} x_i x_j)$$

There is no closed form solution to the score equation in (3.22) and so the MLE for  $\beta$  fails. Estimates for  $\beta$  are therefore obtained numerically using other techniques, such as the Fisher's Algorithm.

Nelder and Weddeburn, (1972) estimated  $\hat{\beta}$  in GLM's using the fisher's scoring algorithm, which is based on the numerically inclined Newton – Raphson method for evaluating maximum likelihoods. For the Poisson regression model with log link,

$$g(\mu) = \log \mu$$

Differentiating the link function with respect to  $\mu$  gives

$$g'(\mu) = 1/\mu$$

The algorithm works by first choosing an initial estimate for  $\beta$ , say  $\beta_0$ . The algorithm updates it to  $\beta_{t+1}$  by successive iterations using

$$\beta_{t+1} = \beta + \left\{ E \left( \frac{\partial^2 \ell}{\partial \beta \partial \beta} \right)^{-1} \right\} \frac{\partial \ell}{\partial \beta}$$

It can be shown for GLM's that the updating equation is

$$E \left( \frac{\partial^2 \ell}{\partial \beta_j \partial \beta_j \dots \partial \beta_j} \right) = - \left( \sum_{i=1}^n x_{ij} x_{il} \dots x_{ir} e^{x_i \beta} \right).$$

Both derivatives (if they exist) are evaluated at  $\beta$ , and the expectation evaluated as if  $\beta$  were the true parameter values.  $\beta$  is then replaced by  $\beta_{t+1}$  and updating is repeated until convergence.

### 3.4 Significance of Regression Parameters

After parameter estimation and fitting of regression models, the next important step is to assess the significance of the individual parameters; the overall model with  $\rho$  coefficients for the predictors included as well as nested models. This can be achieved by using the likelihood ratio which tests the null hypothesis

$$H_0: \beta_1 = \beta_2 = \dots = \beta_\rho = 0$$

The test statistic is

$$G = -2(\text{LL}_c - \text{LL}_\rho)$$

where

$\text{LL}_c$  = log likelihood of the model with the constant coefficient  $\beta_0$

$\text{LL}_\rho$  = log likelihood of the model with all coefficients of parameters present

$G$  has a chi-square distribution with  $P$  degrees of freedom, that is,  $G \sim \chi^2(P)$  and a small  $p$  - value leads to rejecting  $H_0$  and the conclusion that at least one (or more) of the  $P$  coefficients are different from zero (0). A rejection of  $H_0$  leads to a follow-up test to assess the importance of

each explanatory variable in the model, by carrying out statistical test of significance on the individual coefficients.

Testing the significance of individual, and for that matter, single regression parameters in a model can be done using the Wald test or the likelihood ratio test.

### *The Wald Statistic*

The Wald  $\chi^2$  statistic is used to test the significance of individual coefficients in the model by testing the null hypothesis;

$$H_0: \hat{\beta}_j = 0 \text{ (} X_j \text{ is not related to the response)}$$

The Wald Statistics is given by;

$$\left( \frac{\text{Coefficient}}{\text{SEcoefficient}} \right)^2 = \left( \frac{\hat{\beta}_j}{\text{SE}\hat{\beta}_j} \right)^2$$

And this follows the chi- square distribution with one (1) degree of freedom.

That is

$$\text{Wald Statistics} \sim \chi^2_{(1)}$$

It is important to note that for data that produce large estimates of the coefficient, the standard error (SE) is often inflated resulting in lower Ward statistics, and therefore the explanatory variable may be incorrectly assumed to be unimportant in the model. For this reason the likelihood ratio test is generally considered to be more powerful.

### *Likelihood Ratio test (Single Parameter)*

This tests for the significance of a parameter, say  $\hat{\beta}_j$  by comparing the likelihood of obtaining the data when the parameter is zero (denoted by  $L_0$ ) with the likelihood of obtaining the data evaluated at the MLE of the parameters (denoted by  $L_1$ ). It also tests the null hypothesis;

$$H_0: \hat{\beta}_j = 0$$

and the test statistic is given by

$$-2 \ln (\text{Likelihood ratio})$$

$$-2 \ln \left( \frac{L_0}{L_1} \right)$$

$$-2(\ln L_0 - \ln L_1)$$

The test statistics is compared with a  $\chi^2$  distribution with 1 degree of freedom.

### 3.5 Models for correlated data

The choice of models for longitudinal data analysis largely depends on the objective the study seeks to achieve. For a typical longitudinal data it is assumed that measurements from a single subject over a period of time are correlated in some way, unlike GLM's which assumes independence. Because of the presence of correlation as a result of the repeated measures, models under the GLM will be no more applicable, instead we rely on more robust models that can correct for these correlations. Conditional and marginal models are the two most widely used techniques to model non- gaussian (discrete or categorical) responses that are correlated.

#### 3.5.1 Conditional Models

Conditional models assumes that measurements from a single subject taken repeatedly over a period of time share a set of latent, unobserved random effects which can be used to generate an

association structure between the repeated measurements. These random effects are assumed to remain constant within a cluster but changes across clusters. The Generalized linear mixed model (GLMM), which is a conditional model, is the technique used extensively to model such non gaussian correlated responses.

*Definition*

Suppose that given a vector of random effects  $r$ , the responses  $y_1, \dots, y_n$  are (conditionally) independent such that the conditional distribution of  $y_i$  given  $r$  is a member of the exponential family with pdf

$$f_i(y_i|r) = \exp\{a(y)b(\theta) + c(\theta) + d(y)\} \quad (3.23)$$

where

$a(\cdot)$ ,  $b(\cdot)$ ,  $c(\cdot)$  and  $d(\cdot)$  are known functions.

The quantity  $\theta$  is associated with the conditional mean

$$\mu = E(y_i|r)$$

which in turn is associated with a linear predictor

$$\eta = x_i'\beta + z_i'r \quad (3.24)$$

where

$\eta$ : is a random linear predictor

$\beta$ : is a  $(p \times 1)$  vector of fixed regression coefficients

$x_i$ : is a vector of all predictor variables

$z_i$ : is a vector of all variables that have random components

$\gamma_i$ : is a vector of random cluster specific effect

Equation (3.22) above connects through a known link function  $g(\cdot)$  such that

$$g(\mu) = \eta$$

which gives

$$g(\mu) = x_i' \beta + z_i' r, \quad \text{where } \mu = E(Y_i)$$

It is further assumed that the vector of random effects  $r \sim N(O, D)$  where the covariance matrix  $D$ , may depend on a vector  $\alpha$  of unknown variance component.

Although conditional models acknowledge association inherent in longitudinal data, they do not permit generalization of results to the entire population of study.

### 3.5.2 Marginal Models

Marginal models provide a broad class of regression models that are suitable for analyzing diverse types of clustered data in the form of continuous, binary and count data. The main unique feature of the marginal model is that they focus on each response within the cluster separately, allowing dependence of the mean response on covariates, but not on any random effect or other responses within the same cluster. These models focus on the overall population mean response, averaged over all cluster. Hence the regression coefficient relating the mean response to covariates are said to have population averaged interpretations.

In a marginal model, responses  $y_{ij}$  are modeled without explicitly modeling the differences between clusters. The dependency induced by clustered data are dealt with by assuming a form or pattern for the within cluster covariance matrix and this form is assumed to be the same for all

clusters. Distinctively, marginal models are models that are not conditional on unobserved random effect, and hence include population averaged models, as well as models implied by random effects.

Marginal models are appropriate when inferences about the population average are the pointing focus, (Diggle et al, 2002), or when future applications of results require the expectation of the response as a function of the current covariates, (Pepe and Anderson, 1994). These models are best analyzed using the Generalized Estimating Equation (GEE).

### **3.6 The Generalized Estimating Equations (GEE)**

#### **3.6.1 Introduction**

In the 1980's alongside the development of Mixed- effect Random Model (MRMs) and Covariance Pattern Models (CPMs) for incomplete longitudinal data, generalized estimating equation (GEE) models were developed by Liang and Zeger (1986). Basically, GEE is an extension of the GLM for the situation of correlated data. The GEE class of models have become very popular especially for the analysis of categorical and count outcomes, though they work perfectly for continuous outcomes as well. A distinctive feature of the GEE is that it uses quasi – likelihood estimation, which is applied when the specification of the distribution about the data is unknown. Thus the estimation equations are derived without full specification of the joint distribution of a subject's response.

In other words, GEE models avoids the need for multivariate distributions by only assuming a functional form for the marginal distribution at each time point,  $Y_{ij}$ , hence termed marginal models. According to Fitzmaurice et al.,(2004) “the term marginal in this context indicates that

the model for the mean response depends only on the covariates of interest and not on any random effects or previous responses” (Hedeker and Gibbons, 2006).

Furthermore, a basic premise of the GEE approach is that one is primarily interested in the regression parameters  $\beta$  and is not interested in the variance – covariance pattern of the repeated measures. In that regard, the variance – covariance matrix is treated as a nuisance, which must be accounted for in some way in order to perform meaningful statistical tests for the regression parameters.

Also, the GEE models originally developed by Liang and Zeger (1986) allows for separation of the estimating equations for the regression parameters and the association parameters. Thus, in statistical terms these two parameter vectors are assumed to be orthogonal to each other.

Besides subsequently to the development of the GEE models by Liang and Zeger, another class of GEE models were developed, that do not assume orthogonality of these parameter vectors. Due to this development, the original GEE class of models is sometimes referred to as GEE 1 while the latter as GEE 2. GEE 1 is the class of models that is most commonly found in statistical software implementations.

### **3.6.2 GEE Method Outline**

As already stated, the GEE model is an extension of the GLM for situations of correlated data. Thus, the GEE method also shares the properties of the GLM as having three components discussed earlier.

The following represents the major outlines for fitting GEE models.

- (i) Relate the marginal response  $\mu_j = E(y_{ij})$  to a linear combination of the covariates through a link function  $g(\cdot)$  such that

$$g(\mu_j) = X_{ij}\beta$$

where

$y_{ij}$  is the response for subject  $i$  at time  $j$ .

$x_{ij}$  is a  $p \times 1$  vector of covariates

$\beta$  is a  $p \times 1$  vector of unknown regression coefficients

- (ii) Describe the variance of  $y_{ij}$  as a function of the mean,  $V(y_{ij}) = V(\mu_j)\phi$

where

$\phi$  is a possibly unknown scale parameter

$V(\cdot)$  is a known variance function.

- (iii) An important aspect in GEE is specifying the form of the correlation matrix,  $R_i(\alpha)$ .

That is, we choose a form of  $n \times n$  “working” correlation matrix  $R_i$  for each  $y_i$  such

that the  $(jj')$  element of  $R_i$  is the known, hypothesized or estimated correlation

between  $y_{ij}$  and  $y_{ij}$ .  $R_i$  depends on a vector of association parameters denoted by

$\alpha$  which are assumed to be the same for all subjects.

A valuable feature of GEE models is that, it yields consistent and asymptotically normal regression estimates even if this assumed correlation structure is misspecified or incorrect. They represent the average dependence among the repeated observations across subjects.

### 3.6.3 Working Correlation Forms

GEE has a “working” correlation  $\mathbf{R}$  of the repeated measurements.  $\mathbf{R}$  is a matrix of size  $n \times n$  because one assumes that there is a fixed number of time-points  $n$  that subjects are measured at. This matrix is used to represent the association structure, but not the true correlation structure underlying the data set. One desirable property of the GEE is that its method yields consistent estimates of the regression coefficients and their standard errors, even with misspecification of the correlation structure. It is generally recommended that the choice of  $\mathbf{R}$  should be consistent with the observed correlation.

Generally, it is also agreed in principle that an estimator (of some population parameter) based on a sample of size ‘ $n$ ’ will be consistent if its value gets closer and closer to the true value of the parameter as ‘ $n$ ’ increases. So though GEE method yields consistent estimates even with incorrect choice of  $\mathbf{R}$ , efficiency or statistical power is lessened.

However, the loss of efficiency reduces as the number of subjects gets large. It is based on this assertion that O’Muircheartaigh and Francis (1981) find out that loss of efficiency from an incorrect choice of  $\mathbf{R}$  is lessened as the number of subjects gets large.

The commonest forms of working correlations considered are discussed below.

#### *The independence working correlation*

This is the simplest  $n \times n$  identity working correlation matrix of the form

$$\mathbf{R}(\alpha) = \mathbf{I}$$

The independence form is equivalent to assuming no correlation between the repeated measures within clusters. This structure in general may seem not to make logical sense, since data for longitudinal studies are usually highly correlated. Fitzmaurice (1999) reveals the large efficiency loss for longitudinal binary outcomes with time varying covariates using the independent structure. In contrast, Pepe and Anderson (1994) indicates that the use of the independence structure does have certain advantages for models that include time – varying covariates.

The independence structure is given as

$$\begin{pmatrix} 1 & 0 & . & . & . & 0 \\ 0 & 1 & . & . & . & 0 \\ . & . & 1 & . & . & 0 \\ . & . & . & 1 & . & 0 \\ . & . & . & . & 1 & . \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = I$$

*The Exchangeable (compound symmetry) working correlation*

This correlation structure assumes that all of the correlations in  $\mathbf{R}$  are the same. Thus, this structure specifies that  $\mathbf{R}_i(\alpha) = \rho$ . It has the structure shown below.

$$\begin{pmatrix} 1 & \rho & . & . & . & \rho \\ \rho & 1 & . & . & . & \rho \\ \rho & \rho & 1 & . & . & \rho \\ . & . & . & 1 & . & \rho \\ . & . & . & . & 1 & \rho \\ . & . & . & . & . & 1 \end{pmatrix}$$

*The First Order Auto Regressive (AR – 1) correlation structure*

This correlation structure weights the correlation within clusters by their separated time and shows a diminishing correlation coefficient for further distances. It also requires estimating just one parameter similar to the exchangeable. In other words, the within subject correlation over time is an exponential function of the lag namely

$$R_{\tau}(\alpha) = \rho^{|j-i|}$$

Its structure is given as

$$\begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^k \\ \rho & 1 & \rho & \dots & \rho^{k-1} \\ \rho^2 & \rho & 1 & \dots & \rho^{k-2} \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \rho^k & \rho^{k-1} & \rho & \dots & \dots & 1 \end{pmatrix}$$

*The Unstructured/Unspecified working correlation*

With this correlation structure, no constraints are placed on the correlations and we assume different correlations between any two measurements. This form is most efficient, but most useful when there are relatively few time points, as it would be required to estimate all  $n(n-1)/2$  correlations of  $R$ . It becomes however complicated when there are missing data.

Its correlation structure is given as

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \dots & \dots & \rho_{1n} \\ \rho_{21} & 1 & \rho_{23} & \dots & \dots & \rho_{2n} \\ \rho_{31} & \rho_{32} & 1 & \dots & \dots & \rho_{3n} \\ \dots & \dots & \dots & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots & 1 \\ \rho_{n1} & \rho_{n2} & \rho_{n3} & \dots & \dots & 1 \end{pmatrix}$$

### 3.6.4 GEE Estimation

Define  $\mathbf{A}_i$  to be  $n \times n$  diagonal matrix with  $V(\mu_j)$  as the  $j$ th diagonal elements. As already mentioned, we define  $\mathbf{R}_i$  as  $n \times n$  “working” correlation matrix of the  $n$  repeated measures

Then the working variance covariance’s matrix for  $y_i$  is given as

$$V(\alpha) = \phi \mathbf{A}_i^{1/2} \mathbf{R}_i(\alpha) \mathbf{A}_i^{1/2}$$

So for normally distributed outcomes with homogenous variance across time, we have

$$V(\alpha) = \phi \mathbf{R}_i(\alpha)$$

Now, GEE is an extension of the GLM, so from the equation in (3.20), the estimator of  $\beta$  is the solution of

$$\sum_{i=1}^n D_i [V(\alpha)]^{-1} (y_i - \mu) = 0 \tag{3.25}$$

where

$$\hat{\alpha} \text{ is a consistent estimate of } \alpha \text{ and } D_i = \partial \mu / \partial \beta$$

For example, for the normal case,

$$\mu = X_i^* \beta$$

hence

$$D_i = \frac{\partial \mu}{\partial \beta} = X_i$$

and

$$V(\hat{\alpha}) = \phi R(\hat{\alpha})$$

So we have

$$\sum_{i=1}^n x_i' [R(\hat{\alpha})]^{-1} (y_i - x_i \beta) = 0 \quad (3.26)$$

Therefore solving equation (3.26) gives

$$\hat{\beta} = \left[ \sum_{i=1}^n x_i' [R(\hat{\alpha})]^{-1} x_i \right]^{-1} \left[ \sum_{i=1}^n x_i' [R(\hat{\alpha})]^{-1} y_i \right] \quad (3.27)$$

Equation (3.26) is essentially the same as the weighted least squares (WLS) estimator, with the weight matrix being  $[R(\hat{\alpha})]^{-1}$ . A specific limitation is that in WLS, the weight matrix is typically known and however depends on parameters,  $\alpha$ , to be estimated.

In this case, the solution can proceed using Iteratively Reweighted Least Squares (IRLS) where iterative estimates of  $\alpha$  are used to yield new estimates of  $\beta$  with the procedure continuing until convergence (Hedeker and Gibbons, 2006).

From equation (3.26), we notice that the equation depends only on the mean and the variance of  $y$ , hence these solutions yield quasi – likelihood estimates. Therefore GEE solutions iterate between the quasi – likelihood solution for estimating  $\beta$  and a robust method for estimating  $\alpha$  is a function of  $\beta$ .

This basically involves the two steps below, which are repeated until convergence;

- (a) Given estimates of  $R_i(\alpha)$  and  $\phi$ , calculate estimates of  $\beta$  using iteratively reweighted least squares (IRLS).
- (b) Given estimates of  $\beta$ , obtain estimates of  $\alpha$  and  $\phi$ , by calculating Pearson's residuals given by

$$\gamma_{ij} = \frac{y_{ij} - \hat{\mu}_{ij}}{\sqrt{[V(\hat{\alpha})]_{ij}}}$$

and use these residuals to consistently estimate  $\alpha$  and  $\phi$ .

### 3.6.5 Inference

After estimating regression coefficients, we make inferences on them by performing statistical tests of hypothesis and constructing confidence intervals. These would be based on the standard errors associated with the estimated coefficients, obtained by taking the square root of the diagonal elements of the matrix  $V(\hat{\beta})$ .

GEE provides two versions for estimating the matrix  $V(\hat{\beta})$ . (with  $\hat{V}_i$  denoting  $V_i(\hat{\alpha})$ )

- (i) Naive or "Model – based"

Here, the variance of the estimated coefficients is given by

$$V(\hat{\beta}) = \left[ \sum_{i=1}^n D_i \hat{V}_i^{-1} D_i \right]^{-1}$$

where  $D_i$  and  $V_i$  denote their usual meanings

- (iii) Robust or "Empirical/ or "Sandwich" Estimator

Under this approach, the variance for the estimated coefficient is given as

$$V(\hat{\beta}) = M_0^{-1} M_1 M_0^{-1}$$

where

$$M_0 = \sum_{i=1}^N D_i \hat{V}_i^{-1} D_i$$

$$M_1 = \sum_{i=1}^N D_i \hat{V}_i^{-1} (y_i - \hat{\mu})' (y_i - \hat{\mu}) \hat{V}_i^{-1} D_i$$

hence, putting component parts together, we have

$$V(\hat{\beta}) = \left[ \sum_{i=1}^N D_i \hat{V}_i^{-1} D_i \right]^{-1} \left[ \sum_{i=1}^N D_i \hat{V}_i^{-1} (y_i - \hat{\mu})' (y_i - \hat{\mu}) \hat{V}_i^{-1} D_i \right] \times \left[ \sum_{i=1}^N D_i \hat{V}_i^{-1} D_i \right]^{-1}$$

From the above, note that if  $\hat{V}_i = (y_i - \hat{\mu})(y_i - \hat{\mu})'$  then the two estimators are equal, and this only occurs if the true correlation structure is correctly modeled.

Moreover, the robust or “sandwich” estimator, in general, provides a consistent estimate of  $V(\hat{\beta})$  even if the working correlation structure  $R_i(\alpha)$  is not the true correlation of  $y_i$  (Hedeker and Gibbons, 2006).

### 3.6.6 Limitations of the GEE

(i) The assumption can correspond to a mathematically impossible model, in that, some combinations of correlations and variances are mathematically impossible, no matter how the data is generated.

(ii) The choice of working correlation matrix affects the standard errors. A bad choice of

$R$  will generally inflate the standard errors. Inflated standard errors are a serious cost of not modeling the covariance correctly.

- (iii) It does not allow or make individual subject predictions.
- (iv) GEE methods is not good for highly unbalanced data sets, but most suited to balanced longitudinal designs where;
  - number of observation,  $N$ , is large
  - number of repeated measures,  $n$ , is small.

## CHAPTER FOUR

### DATA ANALYSES

#### 4.0 Introduction

This chapter presents a summary results of the analyses of the marginal effects of the school feeding programme, the capitation grants and other observed covariates such as number of teachers (trained and untrained), number of text books per child, location of school, number of classrooms in good condition, gender, availability of portable water and availability of toilet facility on basic school enrolment. This is modeled through the use of Generalized Estimating Equations (GEE) with respect to the four working correlations discussed in chapter three.

#### 4.1 Definition of variables

In order to proceed with the analyses, there is the need to understand the various variables used for the study, and their levels if any. The enrolment data was collected over eight (8) years period from 2008 to 2015 and gathered information on the covariates; number of

trained/untrained teachers in the school, number of classrooms in good condition and number of text books per child.

The categorical variables involved in the study includes; availability of toilet facility (yes or no), gender (males and females), availability of portable water (yes or no), school feeding beneficiary (yes or no), capitation grant beneficiary (yes or no) and location of school (rural, urban or semi-urban).

#### 4.2 Preliminary analyses

This section deals with the exploratory aspect of the analysis, and therefore explores the data set with the help of descriptive measures such as frequency tables and graphs. Descriptive statistics describes the basic features of a data set.

Table 4.2.1: Summary distribution of categorical study variables.

Factor	Number	Percentage (%)
Availability of toilet facility		
Yes	513	79.2%
No	135	20.2%
Availability of portable water		
Yes	384	59.3%
No	264	17.3%
Capitation grant		
Yes	536	82.7%
No	112	17.3%
Location of School		
Urban	240	37.0%
Semi-urban	248	38.3%
Rural	160	24.7%
School feeding		
Yes	115	17.7%
No	533	82.3%

From the Table 4.2.1, we observe that only a small percentage, approximately 18%, of the basic schools in the municipality are beneficiaries of the school feeding programme. Also it can be seen that from the total number of basic schools, a majority of about 83% receives capitation grant from government. We can therefore infer that about 17% of the total number of basic schools who do not receive capitation are private schools, while the rest are public.

Table 4.2.2 Summary distribution of continuous study variables.

Variable	Number	Minimum	Maximum	Mean	Std. Dev.
Number of trained teachers	648	0	15	3.52	2.814
Number of untrained teachers	648	0	10	2.87	2.196
Number of text books per child	648	2	6	4.11	0.911
Number of classrooms in good condition	648	3	6	5.07	1.142
Total Enrolment	648	31	659	214	106

Table 4.2.2 depicts the summary distribution of the continuous variables used in the study. It is observed that the number of trained teachers in the municipality ranges from Zero (0) to fifteen (15) and has the highest standard deviation of 2.814. Also, number of text books per child ranges from two (2) to six (6), and has the least standard deviation of 0.911.

Furthermore, the response variable (total enrolment) recorded a minimum of 31 and a maximum of 659. It has a mean of 214 and a standard deviation of 106 (variance 11236). For Poisson counts, we expect the mean to be equal to the variance, hence the values for the mean and variance in this case suggests over dispersion in total enrolment.

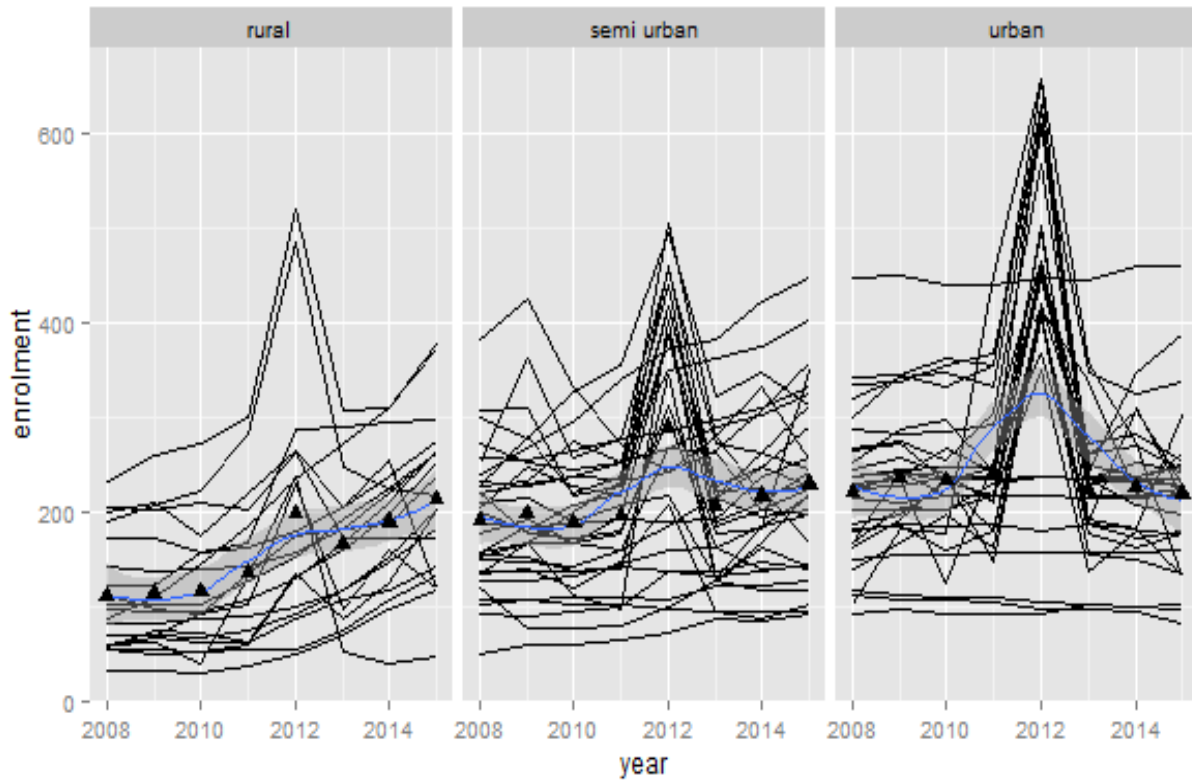


Figure 4.1: Plot of total enrolment by location of school

Figure 4.1 shows a plot of enrolment by years across for the three levels of location; rural, semi urban and urban. It can be observed from the diagram that enrolment over the years increased steadily across the three levels of location. Again, the diagram shows a sharp (steep) increase in enrolment for the period in 2012 for all basic schools in the municipality irrespective of the location.

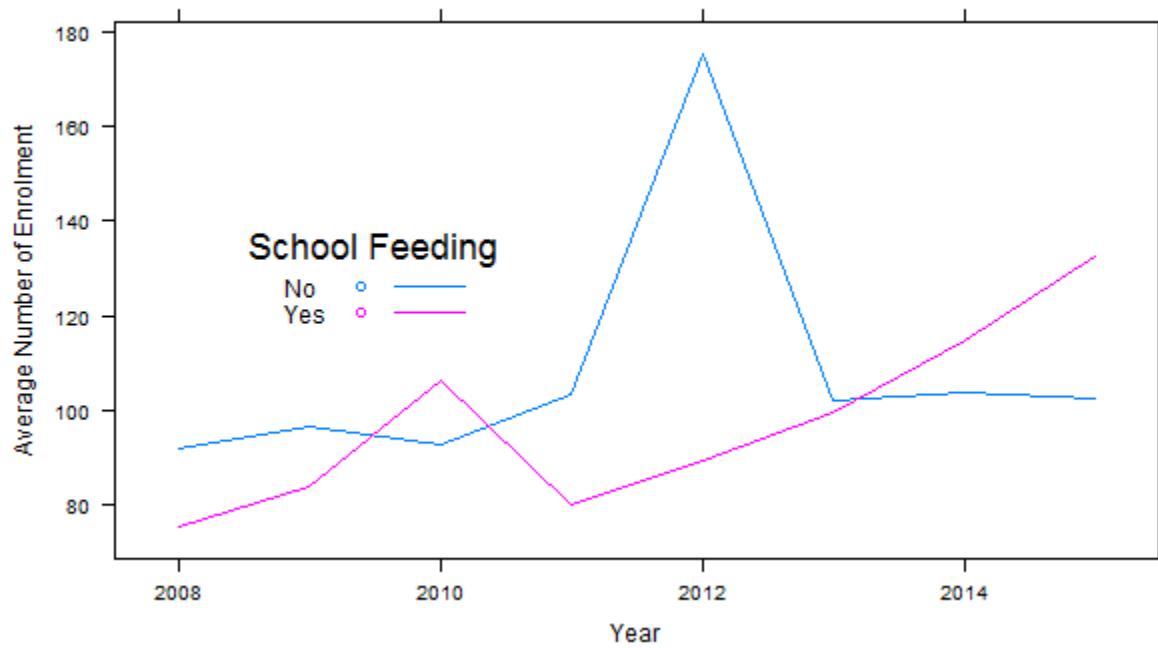


Figure 4.2: Mean Plot for School Feeding

Figure 4.2 depicts the mean number of enrolment over time by school feeding status. From the diagram, it appears schools which enjoy feeding have a fairly increasing trend in enrolment, as opposed to schools without the feeding programme which does appear to show a slight increase in enrolment trend.

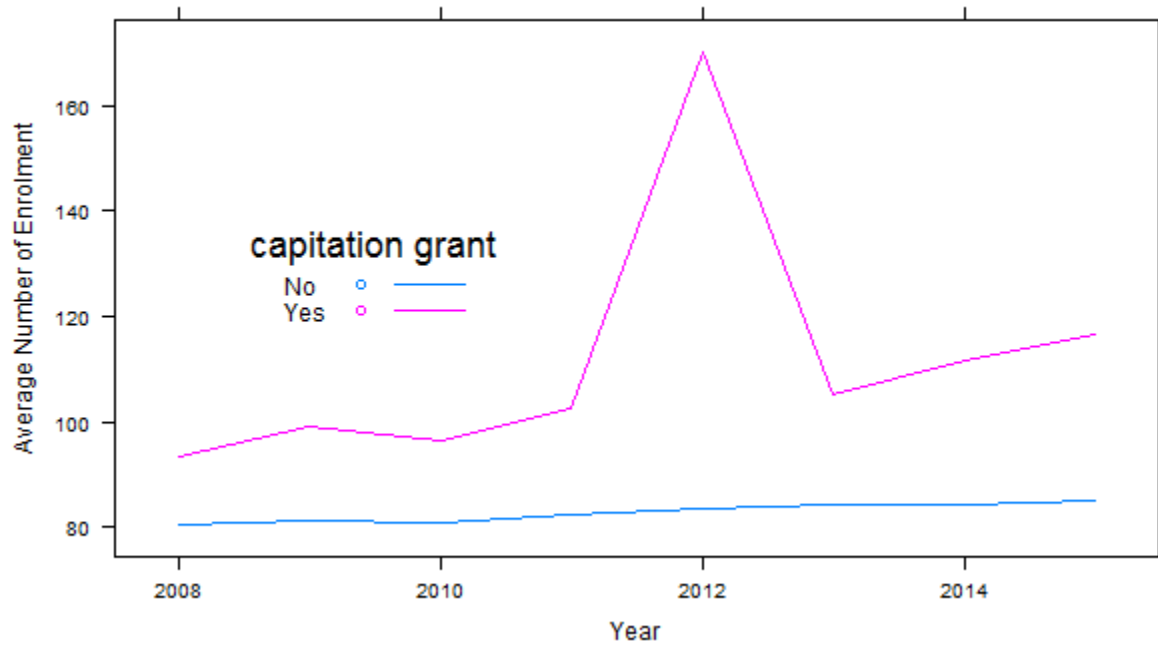


Figure 4.3: Mean Plot for Capitation Grant

Figure 4.3 depicts the mean evolution of number of enrolment by capitation grant. From the diagram, it appears schools which enjoy capitation grant have a fairly increasing trend in enrolment. However, it also appears that schools which do not enjoy the grant have no significant enrolment overtime.

Pearson correlation coefficient matrix

The Pearson’s correlation matrix between enrolments for the period 2008 to 2015 is given below.

1							
0.967	1						
0.935	0.944	1					
0.824	0.854	0.890	1				
0.694	0.757	0.778	0.816	1			
0.677	0.711	0.768	0.794	0.625	1		
0.626	0.640	0.712	0.727	0.559	0.909	1	
0.518	0.498	0.593	0.539	0.363	0.758	0.858	1

The correlation matrix above suggest a generally diminishing correlation coefficient between measurements separated in time

### 4.3 Further Analyses

This section presents the analyses of the GEE family of models which will form the basis of conclusion for the study. The dependent variable, enrolment, is counts and is regressed on a set of predictors. The main objective is to find out the effect of the capitation grant and the school feeding programme on enrolment overtime. Thus, the analysis would include the main effects model together with time interactions with capitation grant and school feeding. This would help us to ascertain if these variables have significant effect on enrolment overtime.

Thus, the analysis begins by considering all the variables as main effects and time interactions with the two main variables under study. It is important here to note that non –significant or non

contributing variables are removed step- wise until a final model of significant variables are obtained. The tables below show the outputs of the GEE models used for the study.

Table 4.3.1: Summary for Poisson GLM (quasi)

Coefficient	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	3.6491	0.1166	31.30	<0.001*
Gender (male)	0.0540	0.0235	2.30	0.0216*
Trained	0.0471	0.0061	7.71	<0.001*
Untrained	0.0363	0.0075	4.87	<0.001*
Sch. Feeding (yes)	-0.4039	0.1050	-3.85	0.0001*
Cap. Grant (yes)	0.1785	0.0811	2.20	0.0280*
Classrooms	0.0482	0.0184	2.62	0.0088*
Portable water (yes)	0.1427	0.0405	3.52	0.0005*
Toilet facility (yes)	0.2547	0.0383	6.65	<0.001*
Time	-0.0084	0.0146	-0.58	0.5642
Sch. Feed (yes)*time	0.0754	0.0172	4.39	<0.001*
Cap. Grant (yes)*time	0.0157	0.0163	0.97	0.3333

*Significant code: '\*' 0.05, Estimated dispersion parameter=19.1*

The Poisson GLM assumes independence among the repeated measures. From Table 4.4.2, we notice that the mean for the response (enrolment) is 214 and the variance being 11236. For Poisson regression model, this wide disparity between the mean and variance constitutes over dispersion which must be estimated. To correct this, the Quasi Poisson GLM is preferred to the

traditional Poisson GLM. The Quasi- Poisson GLM is similar to that of the Poisson GLM, but here takes into account over dispersion.

Table 4.3.1 shows the output for the Quasi Poisson GLM, and it was found out that time interaction with School feeding was significant at the 0.05 level of significance whiles time interaction with capitation grant was not.

#### **4.3.1. Analyses of GEE family of models**

This section of the analyses deals with the estimation of model parameters and its associated standard errors using GEE. The estimation of parameters in GEE is based on the working correlations introduced in the previous chapter. These are the Independence, Exchangeable, Unstructured and AR (1).

#### **4.3.2 GEE “Independent” Model**

Independence assumes that there is no correlation within the responses measured over time and the model becomes equivalent to standard normal regression. The “working” correlation matrix for the independent model was given in chapter three.

The matrix is the estimated covariance matrix used for the parameter estimates for the GEE independent model. From our assumption of the GEE independent working correlation, the dependent variables are uncorrelated and are independent of each other across time (for all the eight years).

Table 4.3.2: GEE ‘Independent’ model

Coefficient	Estimate	Robust Std. Error	Naïve Std. Error	Pr(> z )
(Intercept)	4.1011	0.1335	0.0760	<0.001*
Gender (male)	0.0540	0.0228	0.0239	0.0178*
Trained	0.0471	0.0130	0.0057	0.0003*
Sch. Feeding (yes)	-0.3877	0.1228	0.1065	0.0016*
Cap. Grant (yes)	0.0959	0.1423	0.0801	0.5007
Toilet Facility (yes)	0.3788	0.0954	0.0303	<0.001*
Time	0.0086	0.0056	0.0142	0.1223
Sch. Feed (yes)*time	0.0782	0.0181	0.0174	<0.001*
Cap. Grant (yes)*time	-0.0043	0.0110	0.0157	0.3303

*Significant code: ‘\*’ 0.05, Estimated scale parameters: intercept=19.7, Std. Error=2.58,*

Table 4.3.2 gives the parameter estimates and standard errors for GEE “independent” main and interaction effects. For the main effect model, the intercept, gender effect, number of trained teachers, availability of toilet facility and school feeding were significant at the 0.05 level of significance.

However, for the model with time interactions, only school feeding was found to be the significant variable. The capitation grant interaction with time has no significant effect on enrolment.

### 4.3.3 GEE “Exchangeable” (compound symmetry) Model

The exchangeable working correlation specification allows for constant correlations between any two (2) measurements within a school for all the eight (8) years. With this, we only need to estimate one parameter.

Table 4.3.3: GEE ‘Exchangeable’ model

Coefficient	Estimate	Robust Std. Error	Naïve Std. Error	Pr(> z )
(Intercept)	4.1204	0.1353	0.1195	<0.001*
Gender (male)	0.0540	0.0227	0.0172	0.0180*
Trained	0.0348	0.0078	0.0078	0.0120*
Sch. Feeding (yes)	-0.5268	0.0933	0.0946	<0.001*
Cap. Grant (yes)	0.1463	0.1448	0.1187	0.3130
Toilet Facility (yes)	0.3749	0.0933	0.0935	<0.001*
Time	0.0086	0.0055	0.0102	0.120
Sch. Feed (yes)*time	0.1236	0.0138	0.0137	<0.001*
Cap. Grant (yes)*time	0.0120	0.0106	0.0117	0.0242*

*Significant code: ‘\*’ 0.05, Estimated scale parameters: intercept=20.3, Std. Error=2.59, Estimated correlation (alpha) =0.501*

Table 4.3.3 gives the parameter estimates and standard errors for GEE “Exchangeable” main and interaction effects. Similar to results for the Independence model, gender, number of trained teachers, availability of toilet facility and school feeding as fixed effects, were all significant variables which has an effect on enrolment in basic schools.

More importantly, the school feeding and capitation grant with time interaction was found to have positive effect on enrolment. For the GEE exchangeable working correlation matrix, estimated alpha is 0.501.

#### 4.3.4 GEE AR-1 Model

We now model enrolment and the independent variables using GEE with AR (1) working assumption. Autoregressive GEE model weights the correlation within two years enrolments by their separated time and hence correlation coefficients diminish for further distances. Similar to exchangeable model, it requires only one estimated parameter.

Table 4.3.4: GEE ‘AR-1’ Model

Coefficient	Estimate	Robust Std. Error	Naïve Std. Error	Pr(> z )
(Intercept)	3.9976	0.1859	0.1300	0.1931
Gender (male)	0.0537	0.0232	0.0117	0.0178*
Trained	0.0410	0.0121	0.0098	0.0123*
Sch. Feeding (yes)	-0.469	0.0833	0.1503	<0.001*
Cap. Grant (yes)	0.2498	0.1973	0.1366	0.1376
Toilet Facility (yes)	0.3858	0.0936	0.0583	0.0003*
Time	0.0324	0.0224	0.0223	0.2335
Sch. Feed (yes)*time	0.1040	0.0164	0.0245	<0.001*
Cap. Grant (yes)*time	0.0396	0.0277	0.0247	0.0137*

*Significant code: ‘\*’ 0.05, Estimated scale parameters: intercept=20.2, Std. Error=2.66, Estimated correlation (alpha) =0.873*

Table 4.3.4 gives the parameter estimates and standard errors for GEE “AR-1” main and interaction effects. For the main effect model, gender, number of trained teachers, availability of toilet facility and school feeding were all found to be significant at the 0.05 level of significance.

Again, both school feeding and capitation grant were found to have significant effect on enrolment as years evolve. For the GEE AR-1 working correlation matrix, estimated alpha is 0.873.

#### 4.3.5 GEE Unstructured (Unspecified) Model

For the GEE unstructured working correlation, we assume different correlations between any two years enrolment figures. No constraints are placed on the correlations, which are then estimated from the data.

Table 4.3.5: GEE ‘Unstructured’ Model

Coefficient	Estimate	Robust Std. Error	Naïve Std. Error	Pr(> z )
(Intercept)	3.9922	0.1354	0.1532	<0.001*
Gender (male)	0.2409	0.0267	0.0172	<0.001*
Trained	0.0244	0.0101	0.0089	0.0158*
Sch. Feeding (yes)	-0.4740	0.1234	0.1396	0.0001*
Cap. Grant (yes)	-0.1497	0.1425	0.1437	0.2933
Toilet Facility (yes)	0.4085	0.0982	0.0776	<0.001*
Time	0.0058	0.0067	0.0148	0.3863
Sch. Feed (yes)*time	0.0711	0.0216	0.0182	0.0010*
Cap. Grant (yes)*time	0.0784	0.0117	0.0266	<0.001*

*Significant code: ‘\*’ 0.05, Estimated Scale Paramter:22.93, Estimated Std Error:2.74*

Table 4.3.5 gives the parameter estimates and standard errors for GEE “Unstructured” main and interaction effects. The results for the main effects is not different from the three (3) other GEE models discussed earlier. The intercept, gender effect, number of trained teachers, availability of toilet facility and school feeding were observed as the significant variables at 95% confidence level. It can be also observed that with time interaction, both the feeding programme and the capitation grant have significant effects on enrolment.

Table 4.3.6: Comparison of Naive and Robust Standard Errors

Coefficient	Independence	Exchangeable	AR-1	Unstructured
Intercept	0.0760 (0.1335)*	0.1195 (0.1353)*	0.1300 (0.1859)	0.1532 (0.1354)*
Gender (male)	0.0239 (0.0228)*	0.0172 (0.0227)*	0.0117 (0.0232)*	0.0172 (0.0267)*
Trained	0.0057 (0.0130)*	0.0078 (0.0078)*	0.0098 (0.0121)*	0.0089 (0.0101)*
School Feeding (yes)	0.1065 (0.1228)*	0.0946 (0.0933)*	0.1503 (0.0833)*	0.1396 (0.1234)*
Cap. Grant (yes)	0.0801 (0.1423)	0.1187 (0.1448)	0.1366 (0.1973)	0.1437 (0.1425)
Toilet Facility (yes)	0.0303 (0.0954)*	0.0933 (0.0935)*	0.0583 (0.0936)*	0.0776 (0.0982)*
Time	0.0142 (0.0056)	0.0102 (0.0055)	0.0223 (0.0224)	0.0148 (0.0067)
Sch. Feeding (yes)*time	0.0174 (0.0181)*	0.0138 (0.0137)*	0.0245 (0.0164)*	0.0182 (0.0216)*
Cap. Grant (yes)*time	0.0157 (0.0110)	0.0117 (0.0106)*	0.0247 (0.0277)*	0.0266 (0.0117)*

Significant variables: \*, (.) represents Robust Std. Errors.

Table 4.3.6 shows the standard error estimates (for both naïve and robust) for the four (4) GEE working correlation assumptions used for the study. We observe that the exchangeable assumption gives approximately the same naïve and robust standard error estimates.

#### **4.4 Inference**

Considering the Poisson regression model in tables 4.3.1 we notice that the standard errors are relatively large, which gives an indication that the data cannot be modeled by assuming no correlation among the repeated responses. The failure of the Poisson regression in this case allowed us to use a more robust technique which corrects for the correlations within the clusters, hence GEE.

The choice of GEE models largely depends on the Naïve and Robust standard error estimates. The covariance structure is correctly modeled only when the Naïve standard errors are the same as the Robust. From tables 4.3.6 we observe that the Exchangeable correlation assumption has approximately equal standard error estimates for the two methods. We therefore infer that the underlying association structure can be modeled using the Exchangeable working correlation form.

## CHAPTER FIVE

### DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

#### 5.0 Introduction

The chapter gives a brief report on the discussions, findings and conclusions as a result from the statistical analyses conveyed in the study as well as recommendations based on the findings.

#### 5.1 Discussion

The study sought to assess the effect of the school feeding programme and the capitation grant, among other identifiable variables, on basic school's enrolment in the Wenchi municipality. The other identifiable factors or covariates studied alongside the aforementioned two are gender effect, number of classrooms in good condition, number of textbooks per child, number of trained and untrained teachers a school has, the location of school effect and availability of toilet and portable water .

The study however sought to find out whether these factors have significant impact on enrolment in basic schools. Enrolment figures were collected from 81 basic schools over 8 years period from 2008 to 2015. It is however expected to assume correlation between the enrolment figures within the same school since measurements were taken repeatedly. In recognition of this assumed correlated responses, we employed the GEE family of models to take into account these correlations in the parameter estimation. From the analyses we observe that the estimated parameter coefficients were almost the same under the four working correlation assumption, which reaffirms the property that GEE's produce consistent estimates even if the correlation structure is mis specified.

However, the choice of an association structure largely depends on the estimated standard errors. Another property of GEE is that if the association structure is correctly modeled, then we expect the Naive and Robust standard error estimates to be approximately the same. Table 4.3.6 in Chapter three (3) show a marginally different Naive and Robust standard error estimates for the Independent, Unstructured and AR-1 working correlations. The Exchangeable structure shows approximately similar estimates for the two standard errors. We note the closeness of the standard errors for both empirical and model based in the GEE model for Exchangeable confirming its suitability as a basis for our conclusion.

Thus, the Exchangeable model (Table 4.3.3) shows that school feeding and the capitation grant were significant and has positive effect on basic school enrolment. Also, gender effect was found to be significant, in addition to availability of toilet facility and number of trained teachers in a school. Conversely, number of classrooms, location of school, number of untrained teachers, availability of portable water and number of textbooks were shown to be statistically insignificant.

## **5.2 Conclusions**

Based on the analyses of the data and major findings, we conclude that the school feeding programme and the capitation grant have helped improved enrolment overtime and thus, schools which are beneficiaries of these major educational interventions experienced significant increase in enrolment.

Moreover, the study also revealed that the number of trained teachers in a school has a positive effect on enrolment. Finally, findings of the study revealed that availability of toilet facility is also a factor which affects enrolment in basic schools. Schools with such social amenity will attract more enrolment.

### **5.3 Recommendations**

Based on the above conclusions, we wish to make the following recommendations.

1. The school feeding programme be extended to cover many schools (especially in the semi-urban and rural areas) since it was found to be the major enrolment driver.
2. School authorities should try as much as possible to provide social amenities such as toilet facility and place of urinal in all schools.
3. There is the need to fast track measures aimed at improving girl child education in the municipality.
4. Teacher quality must be given the needed attention. The municipal directorate of education must ensure that adequate trained/qualified teachers are deployed to all schools.
5. Since number of classrooms has no effect on enrolment, there is the possibility of overcrowding, especially in schools which benefit from the feeding programme. This situation negatively affects effective teaching and learning. We therefore appeal to authorities to consider

putting up more classrooms for schools yet to be rolled on to the feeding programme in order to accommodate the corresponding enrolment hikes.

6. We also recommend that further studies be conducted on the effects of the feeding programme and the capitation grant on academic performance.

7. Last but not the least, we recommend that this study be extended to all schools in Ghana, to give it a nationalistic view.

## REFERENCES

- Alaimo, K., Olson, C.M., & Frongillo, E.A. (2001). Food insufficiency and American school-Aged children's cognitive, academic, and psychological development. *Pediatrics*, 108, 44-53.
- Asante, C. (2011). The Capitation grant: Impact on enrolment of pupils in the basic education schools in Ghana. A case study of some selected junior high schools in the Sunyani municipality. KNUST, pp 7-44.
- Asbridge, M., Florence, M., & Veugelers, P. (2008). Diet quality and academic performance. *Journal of School Health*, 78, 209-215.
- Bijlleveld, C.J.H., & Vanderkamp, L.J. (1999). Longitudinal Data Analysis: Designs, Models and Methods. *Sage publications*, pp 241-268.
- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, 101, 147-158.
- Carlin, J. B., Wolfe, R., Coffey, C., & Patton, G. C. (1999). Analysis of binary outcomes in longitudinal studies using weighted estimating equations and discrete-time survival methods: Prevalence and incidence of smoking in an adolescent cohort. *Statistics in Medicine*, 18, 2655-2679.
- Chenoweth, W. (2007). Vitamin B complex deficiency and excess. *Nelson Textbooks of Pediatrics*, 18<sup>th</sup> edition, 271-335.
- Dawuda, A.M. (2012). The Impact of Capitation Grant on Access to Primary Education in Ghana. The Heller school for Social Policy and Management, Brandeis University, pp 6-55.
- Diggle, P. J., Heagerty, P., Liang, K.-Y., & Zeger, S. L. (2002). *Longitudinal data analysis* (2nd ed.). Oxford, UK: Oxford University Press. 434.
- Dobson, J.A. (2002). An Introduction to Generalized Linear Models, Second Edition. Chapman & Hall/CRC texts in Statistical Science Series, Washington D.C., pp 50-58.
- Eysenck, H., & Schoenthaler, S. (1997). Raising IQ level by vitamin and Mineral supplementation: Intelligence, heredity, and environment, *Cambridge University Press*, pp 363-392.
- Fitzmaurice, G.M. (1995). A caveat concerning independence estimating equations with multivariate binary data. *Biometrics*, 51:309-317, 1995.
- Fitzmaurice, G. M., Laird, N. M., & Rotnitzky, A. G. (1993). *Regression models for discrete longitudinal responses*. Statistical Sciences, 8, 284-309.

Frisvold, D. (2012). Nutrition and cognitive achievement: An evaluation of the school breakfast programme, working paper, Emory University.

Ghisletta, D., & Spini-Dunlop, D. (1994). Regression for longitudinal data: A bridge from least squares regression. *The American Statistician*, 48, 299–303.

Hardin, J. W., & Hilbe, J. M. (2002). *Generalized estimating equations*. London, UK: Chapman & Hall/CRC Press.

Heagerty, P. J. (1999). Marginally specified logistic-normal models for longitudinal binary data. *Biometrics*, 55, 688–698.

Hedeker D. (1989). *Random Regression Models with Autocorrelated Errors*. Ph.D. thesis, University of Chicago, Department of Psychology, 1989.

Hedeker, D. & Gibbons, D.R. (2006). *Longitudinal Data Analysis*. Wiley-InterScience, New Jersey, pp131-144.

Horton, N. J., & Lipsitz, S. R. (1999). Review of software to fit Generalized Estimating Equation regression models. *The American Statistician*, 53, 160–169.

Hu, F. B., Goldberg, J., Hedeker, D., Flay, B. R., & Pentz, M. A. (1998). Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *American Journal of Epidemiology*, 147, 694–703.

Jones, T., Borg, W., Boulware, S., McCarthy, G., Sherwin, R., Tamborlane, W. (1995). Enhanced adrenomedullary response and increased susceptibility to neuroglycopenia: Mechanisms underlying the adverse effect of sugar ingestion in children. *Journal of Pediatrics*, 126, 171–177.

Kleinman, R., Murphy, J., Little, M., Pagano, M., Wehler, C., Regal, K., & Jellinek, M. (1998). Hunger in children in the United States: Potential behavioral and emotional correlates. *Pediatrics*, 101 (1), e3.

Khuri, I.A. (2010). *Linear Model Methodology*. Chapman & Hall/CRC texts in Statistical Science Series, Washington D.C., pp 473-479.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38, 963–974.

Liang, K.Y., & Zeger, S. L. (1986). Longitudinal data analysis using Generalized Linear Models. *Biometrika*, 73, 13–22.

- Lipsitz, S. H., Fitzmaurice, G. M., Orav, E. J., & Laird, N. M. (1994). Performance of Generalized Estimating Equations in practical situations. *Biometrics*, 50, 270–278.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models*. 2nd edition, Chapman and Hall, New York, 1989.
- Michikazu, N., & Weiming, K. (2009). Statistical Models for Longitudinal Data. *Applied Mathematical Sciences*, Vol 3, 534-593.
- Molengerghs, G. & Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. Springer Science + Business Media, Inc., New York, pp 151-240.
- Myers, A., Sampson, A., Wietzman, M., Rogers, B., & Kayne, H. (1989). School breakfast programme and school performance. *American Journal of Disease of Children*, 143, 1234-1239.
- Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, Series A*, 135, 370–384.
- Neuhaus, J. M., Kalbfleisch, J. D., & Hauck, W. W. (1991). A comparison of cluster specific and population-averaged approaches for analyzing correlated binary data. *International Statistical Review*, 59, 25–35.
- Norton, E. C., Bieler, G. S., Ennett, S. T., & Zarkin, G. A. (1996). Analysis of prevention program effectiveness with clustered data using Generalized Estimating Equations. *Journal of Consulting and Clinical Psychology*, 64, 919–926.
- Osei, D.R., Owusu, A.G., Asem, E.F., & Afutu-Kotey, R.L. (2009). Effects of Capitation grant on education outcomes in Ghana. *Institute of Statistical Social and Economic Research*, pp 6-20.
- Osei-Fosu, A.K. (2011). Evaluating the impact of the Capitation grant and the School Feeding programme on enrolment, attendance and retention in schools: The case of Weweso circuit. *Journal of Science and Technology*, vol.31, No.1 (2011), pp 55.
- Owusu, D.I. (2011). Application of Generalized Estimating Equation (GEE) model on student academic performance: *A case study of final year Math four students at KNUST*. pp 17-66.
- Pendergast, J. F., Gange, S. J., Newton, M. A., Lindstrom, M. J., Palta, M., & Fisher, M. R. (1996). A survey of methods for analyzing clustered binary response data. *International Statistical Review*, 64, 89–118.

- Pepe, M. S., & Anderson, G. L. (1994). A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. *Communications in Statistics*, 23, 939–951.
- Pollitt, E. (1993). Iron deficiency and cognitive function. *Annual Review of Nutrition*, 13, 521-537.
- Prentice, R. L. (1988). Correlated binary regression with covariates specific to each binary observation. *Biometrics*, 44, 1033–1048.
- Qu, A., Lindsay, B. G., & Li, B. (2000). Improving generalised estimating equations using quadratic inference functions. *Biometrika*, 87, 823–836.
- Redden, J., Wahlstrom, K., & Reicks, M. (2002). Children's perceived benefits and barriers in relation to eating breakfast in schools with or without universal school breakfast. *Journal of Nutrition Education & Behavior*, 34(1), 47-52.
- Sorhaindo, A., & Feistein, L. (2006). The Relationship between child and school outcomes. *Wider benefits of learning research report*. 18, 77-91.
- Wahlstrom, K.L., & Begalle, M.S., (1999). *Nutrition and students academic performance*. Wilder Research publications, 15, 17-29.
- Wang, Y.-G., & Carey, V. (2003). Working correlation structure misspecification, estimation and covariate design: Implications for generalised estimating equations performance. *Biometrika*, 90, 29–41.
- Ware, J. H. (1985). Linear models for the analysis of several measurements in longitudinal studies. *The American Statistician*, 39, 95–101.
- Ware, J. H., Lipsitz, S. H., & Speizer, F. E. (1988). Issues in the analysis of repeated categorical outcomes. *Statistics in Medicine*, 7, 95–108.
- Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gaussian–Newton method. *Biometrika*, 61, 439–447.
- Zajonca, C. (1976). *Generalized Estimating Equation Models for Correlated Data: A Review with Applications*. *American Journal of Political Science*, 45, 470–490.
- Zeger, S. L., & Liang, K.-Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 42, 121–130.

Zeger, S. L., & Liang K.-Y. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73:13-22.

Zeger, S.L., Liang, K-Y. & Albert, S.P. (1988). Models for Longitudinal Data: A Generalized Estimating Equation Approach, *Biometrics*, 44, 1049-1060.

Zheng, B. (2000). *Summarizing the goodness of fit on generalized linear models for longitudinal data*. Statistics inc.

Zorn, C. J. W. (2001). Generalized Estimating Equation models for correlated data: A review with applications. *American Journal of Political Science*, 45, 470–490.

Zuur, A.F., Ienu, N.E., Walker, J.N., Saveliev, A.A., & Smith M.G. (2009). *Mixed Effect Models and Extentions in Ecology with R*. Statistics for Biology and Health, pp 295-320.