UNIVERSITY FOR DEVELOPMENT STUDIES, NAVRONGO

MULTILEVEL MODELLING WITH APPLICATION TO CHILD DISCIPLINE PRACTICES IN GHANAIAN HOMES

STEPHEN KWAKU AMOAH

2020

UNIVERSITY FOR DEVELOPMENT STUDIES, NAVRONGO

MULTILEVEL MODELLING WITH APPLICATION TO CHILD DISCIPLINE PRACTICES IN GHANAIAN HOMES

BY

STEPHEN KWAKU AMOAH (M.Sc., B.Sc. Statistics)

ID NO: UDS/DAS/0014/12

Thesis submitted to the Department of Statistics, Faculty of Mathematical Sciences, University for Development Studies in partial fulfillment of the requirements for the award of DOCTOR OF PHILOSOPHY in Applied Statistics

MAY, 2020

DECLARATION

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

Candidate's Signature…………………………………………………………………..

Date………………………………………………………………………………………..

Name: Stephen Kwaku Amoah

Supervisors'

I hereby declare that the preparation and presentation of this thesis was supervised in accordance with the guidelines on supervision of dissertation/thesis laid down by the University for Development Studies.

Principal Supervisor's

Signature:…………………………………….Date…………………………

Name: Dr. Ezekiel Nii Nortey

Co-Supervisor's

Signature:……………………………………..Date……………………

Name: Dr. Abukari Alhassan

ABSTRACT

This study applied multilevel modelling to analyze data on child discipline practices from 8,156 households from the MICS 4 data collected by GSS in 2011. The analysis took into consideration intra-cluster correlation that results from the collection of such hierarchical structured data and unmeasured higher level characteristics that impact on values of the response variables. A key purpose of the study was to address the error terms of individual observations that correlate in such datasets and lead to the violation of standard statistical assumption of different observations having independent residuals. Autocorrelation which leads to underestimation of standard errors of regression coefficients and results in overestimation of *t* values and subsequent underestimation of *p* values was addressed. The results showed significant contextual effects on household responses at the regional level but not at the location (rural/urban) level. Also, younger household heads (15-21 years), single child households and wealthy households tended not to approve the use of physical discipline for correction. Ethnicity had significant effect on household's approval of physical discipline. In actual application of the different discipline methods by households, the results further showed that, the number of children aged 2-14 in a household, the religion and education level of household heads had significant influence on a household's likelihood to apply physical discipline methods or non-physical discipline methods or psychological aggression discipline methods relative to the reference group of random use of discipline methods. It is recommended that to totally eliminate violence against children we will need to take into consideration information provided in this study to provide parents with nonviolent alternative discipline methods.

ACKNOWLEDGEMENT

This thesis has been possible because of the strong support and contributions from many people. My utmost gratitude goes to my principal supervisor Dr. Ezekiel Nortey for his encouragement and insightful suggestions which enriched this thesis. I am also grateful to my co- supervisor Dr. Abukari Alhassan whose contributions greatly challenged me to improve clarity in the interpretation of the findings. I am most grateful to Prof. N.N.N Nsowah-Nuamah who provided various leads and suggestions. I am also grateful to Dr Sarpong (University for Development Studies) and Dr. Samuel Iddi (University of Ghana) for their useful comments and suggestions.

To my family members whose acceptance of my long working and study hours at the expense of my share of duties at home, I say I am very grateful for your patience without which it would have been difficult to successfully complete this thesis. To all my colleges who urged me on in difficult times, many thanks.

DEDICATION

To my Uncle James Kwesi Kyeremah Appiah, aka ‘Grandpa’

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AIC	Akaike Information Criterion
AICC	Corrected Akaike Information Criterion
ANOVA	Analysis Of Variance
AUC	Area Under Curve
BCG	Bacillus Calmette-Guerin vaccine
BIC	Bayesian Information Criterion
CAIC	Consistent Akaike Information Criterion
CD	Physical Discipline
CDF	Cumulative Distribution Function
DHS	Demographic and Health Survey
DOVVSU	Domestic Violence and Victim Support Unit
DPT	Diphtheria, Pertussis, and Tetanus vaccine
GEEs	Generalized Estimating Equations
GLMM	Generalized Linear Mixed Models
GLS	Generalized Least Squares
GSS	Ghana Statistical Service
HQIC	Hannan-Quinn Information Criterion
ICC	Intra Cluster Correlation
LMM	Linear Mixed Model
LR	Likelihood Ratio
MEM	Mixed Effects Model
MICS	Multiple Indicator Cluster Survey
MMR	Measles, Mumps, and Rubella vaccine

MoGCSP	Ministry Of Gender, Children And Social Protection
OLS	Ordinary Least Squares
PHC	Population And Housing Census
PSU	Primary Sampling Unit
ROC	Receiver Operating Characteristics
SAS	Statistical Analysis System
UNICEF	United Nations Children's Emergency Fund
VPC	Variance Partition Component
-2LL	-2Log Likelihood
MLE	Maximum Likelihood Estimate
GLMM	Generalized Linear Mixed Model

PREFACE

The author of this thesis has also contributed to the writing of the following papers.

1. Amoah, S.K., Nortey, E.N.N, Alhassan, A. (2020). Correlates of Parental Choice of Child Discipline Methods in Ghana: A Multilevel Modelling Approach. Open Journal of Applied Sciences, 10, 78-99. https://doi.org/10.4236/ojapps.2020.103007

2. Amoah, S.K., Nortey, E. & Alhassan, A. Child discipline practices in Ghanaian homes. *J Public Health (Berl.)* (2019). From Theory to Practice. ISSN 2198-1833
 https://doi.org/1007/s10389-019-01079-z

CHAPTER 1

GENERAL INTRODUCTION

1.1 Background to the Study

One of the most studied and researched topics in developmental psychology is discipline and punishment of children by parents and caregivers as a way of teaching them skills and values to become capable adults. When parents and caregivers consider a particular child behaviour as inappropriate they apply discipline with the idea of discouraging that behaviour from being repeated in the future. The various types of disciplinary methods used by parents and caregivers have been associated to children's outcomes (Holden, 2015a). Research has shown that the wrong application of these discipline methods has grave consequences on the later life of the receiving child. Therefore, it is critical to conduct an investigation into the situation in order to come out with policies that will facilitate the eradication of violence against children. This study's concern is modelling the various forms of discipline methods used in Ghanaian homes and their associated risk factors in terms of child upbringing beliefs and practices across the country and help develop effective policy programs to guide child discipline.

Despite the fact that data abounds in child discipline studies, no data have been analyzed using statistical models that takes into consideration the clustering/hierarchical nature of the data collected. The goal of the study therefore, is to find important risk factors linked with child discipline practices and to make recommendations that can reduced violence against children in the Ghanaian society. To achieve this goal, the MICS 4 survey data was reanalyzed to develop a model that can detect the important risk factors linked with child discipline practices.

In statistical modelling, the type of data collected has a critical role in determining the best statistical approach to take. Many kinds of research designs and resulting data structures in social science studies involve multilevel or hierarchical data structures where observations are clustered within groups and the groups in turn are nested in upper groups. "Multilevel modelling is an effective approach for studying the relationship between individuals and their various groups. This is because it allows the incorporation of substantive theory about individual and group processes in the sampling schemes of many research studies (e.g multistage stratified samples, repeated measures designs) or in the hierarchical data structures found in many existing data sets encountered in social science, management, and health-related research" (Heck & Thomas, 2000).

Hierarchical data structures can be categorized into distinct groups or "clusters" in a study. Such grouped data results in cluster-correlated data and can be grouped in a variety of different ways. For example, in educational studies the outcomes, assessments of quality of teaching and learning are often obtained from teachers who are nested within different schools, classrooms and characteristics of their pupils.

Nesting or clustering in data can result from a naturally occurring hierarchy in the target population or a consequence of the study design or both. Naturally occurring hierarchical data structures can have more than two levels e.g. children (level 1 units) nested within classroom (level 2 units), nested within schools (level 3 units).

Data on child discipline practices, like most social science data, often times are made up of complex data structures where data are obtained from complex research designs from populations which exhibit hierarchical groupings or have multilevel structures. With such data, observation at the individual level nested within a higher level are found in natural hierarchies. When data from such populations are modelled it is important to take account of the structures (group effects, as well as individual effects and also interaction between them (cross-level effects) if the data values are related to them.

Such data structures are subject to intra-class correlation where households within the same group or enumeration area are more alike in terms of attitudes and behaviours than in households across groups. In such instances, households within the same unit are affected by the same aggregate effects and thereby lead to correlations within the households. The within group correlation is usually termed group dependence. Jones and Duncan (1995), found that people with nearly identical and socio-economic characteristics but who live in different areas can have divergent health conditions. For example, in geographical studies on health, the local context where individuals live has been found to make a difference with respect to a wide range of people's health outcomes.

"The existence of group dependence among the lower and other higher level units violates the classic assumption of independence in a standard regression analysis, raising the risk of inefficient model estimation and incorrect inference" (Goldstein, 2003a). In such situations, nested hierarchies are created by the different levels and these make the application of traditional logistic regression inappropriate as the variance of the residual errors will be correlated between individual observations. Regression approaches to modelling such data structures that comprise of individual and group effects include multilevel modelling.

Researchers often wish to find what inform or influence people to act or behave in the way they do. To achieve this, most researchers use regression analysis of harmonized data where individual-level outcomes are modelled as a function of both individual and location-level characteristics. In this study, multilevel binary and multinomial sampling-based approaches were extend to model dependent binary and multinomial outcomes with hierarchical structures. The impact of observed covariates on the outcome of interest was investigated by focusing on the precision of the estimators and standard errors, to provide unbiased hypothesis tests.

The study also demonstrates how individual household characteristics and contextual effects can simultaneously be modelled if correlation between the covariates and the error term is suspected in the outcome under study. Multilevel modelling therefore, provide strong designs for observing both contextual and aggregate outcomes these effects produce.

To assess the performance of the proposed models, a series of simulations and diagnostics were conducted to approximate the output of the developed models through repetitive random application of the models' algorithm. The analysis revealed the importance of taking data clustering into account when it exists and proposes appropriate statistical methods that factors-in clustering when they exist in a dataset.

1.2 Problem Statement

Social science research often involves the examination of effects of social context with hierarchical models where individuals are nested in social contexts like schools and neighbourhoods whose effects are believed to shape individuals outcomes. Not factoring in these group contexts in ones attempt to understand individuals' behaviour or attitudes, can seriously limit a researcher's quest to understand underlying structures of interest.

Many national surveys use complex survey design (multi-stage sampling). These survey designs result in data which are clustered with hierarchical structure in the target population. For example in the dataset for MICS 4, households are the level 1 units, area segments or enumeration areas, the level 2 units, and regions the level 3 units. One important consequence of clustering is that measurement on units within a cluster are more similar (i.e. correlated) than measurements on units in different clusters. Therefore, statistical analysis of such data, should take into consideration the intra-cluster correlation that results at each level as failure can lead to erroneous conclusions concerning the impact of the diverse sources of influence on the outcome variable.

It is important to note that higher level units have similar impact on lower level behaviours. For example, in a community with the same health centre available to every community member, every characteristic of this health centre will impact similarly on all who attend the health centre. Likewise, if there are unobserved community factors that have influence on the behaviours of individuals in each community, it will lead to a situation where within each community the individual level outcomes will be correlated after controlling for observed individual-level and community-level predictors.

The presence of unobserved higher level characteristics requires the adjustment of one's analysis to factor-in such structures. This is because their impact on the outcome of interest cannot be controlled and therefore, they are included in the "error term" in statistical models. This implies, the error terms for individual responses in each community could be correlated. This follows from the fact that the unobserved community characteristic that influences a community member to have a high value for the outcome of interest would likely result in other members of the same community having similarly high values of the outcome of interest. The presence of such error correlations among individuals within communities implies that the standard statistical assumption that different observations have independent residuals will be violated. This prompts consideration of important statistical concerns.

Firstly, when unobserved higher-level factors influence behaviours of individuals within locations, statistical properties of estimators which rely on the assumptions of independent residuals, are violated. For instance, the Gauss-Markov Theorem which states that "the ordinary least squares (OLS) estimator is the best linear estimator within the class of unbiased estimators" becomes flawed as it only holds if there is no correlation between the residuals and the observed values. Secondly, the Multilevel error components estimator provides better estimates, as in the case of generalized least squares or maximum likelihood estimates, than ordinary least squares in situations where there are unobserved higher-level effects. Thirdly, multilevel modelling provides correct standard errors and thus correct confidence intervals and significance tests.

Similarity, within a cluster also implies that, to some extent, one can predict the outcome of an observation if one knows the outcome of other observations in the same cluster. This suggests that not every observation provides an independent piece of information and that the total amount of information contained in a sample with clustering is less than that in a sample without clustering. For instance using linear regression techniques and ignoring the fact that the students are grouped into teachers and classes, Bennett (1976), showed that in Great Britain elementary school students benefitted more from a formal style of teaching. These results were widely known and became quite influential until Aitkin et al. (1981), demonstrated that, once the grouping of the students is taken into consideration in a multilevel model, the results obtained by Bennett concerning teaching styles were no longer statistically significant.

When all variations at levels higher than the first level are captured by observed variables, multilevel data can be analysed by traditional linear or nonlinear models. In such a case, conditional on the observed variables, the observations in the same cluster are no longer dependent, and the standard errors obtained by traditional models are correct. On the other hand, if one is dealing with a dataset with a multilevel error structure and does not account for uncaptured higher level variations, it will lead to a situation where you can overstate the importance of the estimated statistical relationships and "uncover" statistically significant relationships when they do not exist. Again, the efficient estimation of the impact of covariates on the outcome of interest will be lost. So, it is importance that one understands the structure of the data one is using as it is critical in determining the statistical techniques to be used for analysis and interpretation.

Unfortunately, the analysis of data on most child discipline studies in Ghana (Section 2.9) have been done only descriptively to establish associations between parenting styles and disciplinary practices without considering the structure of the collected data. Such analysis do not provide in-depth information of the situation at hand as they primarily focused on describing the nature of the phenomenon without focusing on "why" the phenomenon occurs. Though for some of these studies, complex survey design methods were employed (multistage cluster sampling for MICS 3 and 4). These descriptive results may indicate certain level of bias due to the absence of statistical tests and could lead to misleading conclusions and misdirect policy to addressing child discipline issues.

"Parental discipline response often occurs as a result of a complex interplay of factors embedded within a wide system of relationships within the family as well as social economic influences which obstruct or facilitate effective parenting" (Smith et al., 2005). "This obvious hierarchical structure has oftentimes been overlooked leading to substantial errors in interpreting the results of statistical significance tests" (Goldstein, 2003).

To bridge these gaps indicated above, this study sought to conduct a multilevel modelling using the MICS 4 dataset to investigate the factors that influence child discipline practices in Ghanaian homes.

In conclusion, not taking into account obvious structures or clusters in a dataset can lead to incorrect inferences because observations or values then are no longer independent and errors will be correlated within the clusters or groups and lead to autocorrelation within groups. The consequence of autocorrelation leads to underestimation of the standard errors of regression coefficients. When this happens the *t* values become overestimated than they actually are and also leads to underestimating of the *p* values. When this happens it is more likely to reject the correct null hypothesis and lead to invalid conclusions. That is, you would tend to get more significant effects which are not really significant.

1.3 Research questions

To address the research problem, eight research questions were posed to guide the study.

1. What is the probability that a household in Ghana will approve the use of physical discipline as appropriate for child upbringing?
2. Are there significant variability in responses of households across the different locations and regions of Ghana in their decision to approve the use of physical discipline?
3. Do household characteristics influence their decision to approve the use of physical discipline as appropriate for child upbringing?
4. What is the probability of choosing particular discipline methods by household heads?
5. Do the types of discipline methods chosen by household heads vary across the different regions of Ghana?
6. Does the choice of a discipline method by a household dependent on the research predictors?
7. Do specific predictors have influence on households' choice of particular discipline methods?
8. Are there regional effects on a household's decision to choose a particular type of discipline method?

1.4 Objectives of the Study

The main objective of this study is to conduct an in-depth study on child discipline practices in Ghana using a multilevel modelling approach with application to the MICS 4 dataset to determine what household characteristics and contextual factors that influence child discipline practices in Ghana. The study will measure the strength of relationships between the dependent and independent variables by studying the variation in the independent variables and observing what effects they have on the dependent variables. To achieve this main objective, the following two specific objectives were addressed.

1. To examine the factors that influence Ghanaian household heads to approve the use of physical discipline as an appropriate way of bringing up children. This objective addresses the binary response variable in the study dataset by answering research questions 1 to 3.
2. To investigate what causes Ghanaian household heads to choose particular discipline methods to discipline their children. .This objective addresses the multinomial response variable of the study dataset by answering research questions 4 to 8.

1.5 Significance of the Study

Until in 2006 and 2011, when the Ghana Statistical Service, conducted Multiple Indicator Cluster Surveys (MICS) 3 and 4 limited studies on child discipline in Ghana focused on children in schools. The MICS is a nationally representative internationally comparable household survey implemented to examine protective and risk factors of child development in developing countries around the world. The surveys described the situation of children in Ghana in terms of nutrition, parenting, discipline and violence in the home environment. Descriptive analysis of the module on child discipline in the two surveys indicated a high percentage use of physical punishment (corporal punishment). Again, the surveys indicated a worrying trend of increased rise in the application of physical discipline methods to children. In Africa, Ghana was the first country to endorse the 1989 UN convention on the children's rights' (Apt et al., 2012). Yet it has been argued that certain social facts such as poverty and the image of childhood in Ghana make the ratification of the UN convention more as rhetoric than a reality.

The justification of this study was based on the following grounds: Firstly, "Ghana has laws for the protection of children, however, enforcement remains weak. Violence and abuse of children, including sexual abuse, remains very high with over 90 % of children reporting having experienced physical violence, both at home and in the school environment" (Ghana News Agency, 2014).

Child labour and child trafficking continue to be a challenge despite Government and civil society's effort to address them. "More than 4,000 children live in residential homes, often labelled as 'orphanages'. Many of these children are unnecessarily separated from their families. Ghana also has a very high rate of adoption of children, including inter-country adoption. The absence of transparency and control in these adoptions have led to the Government of Ghana placing a moratorium banning all adoptions of children until the situation can be examined further" (Ghana News Agency, 2014).

UNICEF (2014) reported that many children in Ghana were being subjected to physical, emotional and sexual abuse. The specific forms of abuse included defilement, harmful corporal punishment, abandonment and female genital mutilation in some traditional areas. To date, many Ghanaian children are engaged in voluntary and forced labour, and thus denied opportunity to have education.

Secondly, review of limited literature on child disciplinary methods in Ghana show that earlier studies focused mainly on the use of corporal punishment with little or no mention of other methods. In order to understand parental discipline there is the need to expand our focus to the full range of parent's disciplinary practices including psychological and non-physical methods. Furthermore, these earlier studies were based on small samples using the school situation. Again, analysis of these earlier studies were mainly descriptive, did not consider the presence of unobserved or unmeasured higher level or macro-level characteristics of where households are located and did not provide reasons for the observations made.

Thirdly, this study went beyond descriptive findings to provide in-depth secondary quantitative analysis to advance our understanding of child disciplinary methods in Ghana and explain the underlying process linking households and disciplinary methods based on their circumstances. This study provides further understanding of the prevalence and nature of child disciplinary methods and establishes evidence to enrich programme and policy development work on child discipline in Ghana.

Again, the study's findings will be a critical source of information to provide understanding of social attitudes and the risk factors to benefit organizations or practitioners such as UNICEF, Domestic Violence and Victim Support Unit (DOVVSU), the Ministry of Gender, Children and Social Protection (MoGCSP and clinicians for consideration in counseling clients and developing coping strategies. Furthermore, the in-depth information derived from the study could provide understanding of social attitudes and risk factors, to discourage the use and social acceptance of violence against children in our schools and homes. Furthermore, it will also contribute to understanding of the 'drivers' and 'barriers' to change in child disciplinary practices and strengthen the overall child protection system of Ghana. Again, it will strengthen child advocacy efforts and direct future targeted research in child disciplinary methods by facilitating their work in terms of policy decisions, strategy formation and implementation of issues on children. It will also provide valuable information to parents in terms of the connection between parenting styles and its later effects on children.

Fourthly, the study sought to focus on modelling and knowledge discovery to provide statistical generalization of the findings to the population (Neuman, 2000; Yin, 1994; Zikmund, 2000), rather than purely descriptive as has been the case in past studies of child discipline in Ghana.

Finally, with the current interest in ways of engaging children in cooperative interactions rather than control-based ones, under the label of "positive discipline" instead of use of physical discipline and its after effects on children (Holden, 2012), the motivation to develop these models is to use them to draw inferences to understand the relationships between household responses and their characteristics, predict the behaviour of households given their situation and characteristics and to contribute to efforts needed to fully protect children from all forms of violent discipline and them achieving their developmental potential and wellbeing.

1.6 Limitations of the Study

The interpretation of this study should be done with the following limitations in mind**.** Firstly the study used a self-reported measure which has the disadvantage of bias errors when respondents over-report or under-report the assessment of private thoughts, feelings and behaviours in hypothetical situations (Heppner et al., 2008). According to Latikin et al., (2017) “bias errors, results from participants desire to provide socially acceptable or socially desirable responses”. “Given this, it is possible that participants may respond to items on an instrument in a socially acceptable manner, leading to bias errors” he concluded. For instance, in order not to be seen as unreasonable persons for using hash discipline methods such as extreme corporal punishments and for fear of being frowned upon and even to prevent legal retribution, may conceal their true disciplinary practices.

In multilevel analysis the characteristics of each level provides specific information of the effect of that level apart from the general effect of the level. Multilevel models however, are not able to pin point the exact spot in the group or level causing the variation for the group.

Additional information on the characteristics of the selected child (is child the biological child of household head) and a question on whether household head considers the gender of the child before administering discipline. Another useful information missing is whether the selected child has ever lived with other families (ie extended family) other than the current household. Again, the age range of the selected child (2-14) was too wide. The range could have been broken into two groups (1-5) and (6-14). The wide range denied information on how little children (1-5) are discipline compared to children aged 6-14. Do parents take into consideration the age of a child before administering discipline? The inclusion of such data in the study would have thrown more light on the issues at stake and provided a deeper understanding of child discipline practices in the country. Furthermore, the absence of specific information on each region in the dataset made it impossible to determine what in the regions actually caused households within them to response the way they did.

1.7 Organisation of the Dissertation

The remainder of this dissertation is organized as follows: Chapter 2, is a review of the relevant literature on multilevel models, discreet choice responses, and child discipline practices introduced above. Also discussed are types of response outcomes and simulation methods in the context of multilevel /hierarchical data structures. Child discipline concepts and methods used in their analysis are discussed in this chapter. Chapter 3 explains the methodology applied in this study to achieve the research objectives. Specifically, it covers sections on description of the dataset, operational definitions, and models to answer the research questions. Also, principles of multilevel modelling, generalize linear mix models, interpretation of coefficients, model evaluation and diagnostics are covered.

Chapter 4 presents the detailed results, interpretation of estimated parameters and discussions in terms of the study's objectives and key responses. Also presented in this chapter are the evaluation results of the developed models and discussion of their robustness.

Chapter 5 presents the general conclusion and consolidates the answers to the research questions and objectives. It also provides, recommendation, implication of the results and contribution to knowledge.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter discusses the theoretical framework used to conceptualize this study and a review of the relevant literature in relation to the objective of the research. It has two main parts. The first part reviewed literature on hierarchical models and their application, modelling hierarchical/clustered data, approaches to dealing with dichotomous data and simulating hierarchical categorical data to provide the framework of the research.

The second part discussed studies on child discipline methods in terms of theoretical contributions made to it by various schools to thought and application of different types of discipline and probable future effects on the recipient. This part also discussed effects of socio-economic status of families and contextual factors as they influence parenting, family dynamics and children's developmental outcomes. Also, child disciplinary practices in developing countries including Ghana and the various methodologies and models developed in the study of child disciplinary methods were discussed.

The foundation of all social research is statistical analysis and thus the approach or methodology applied is very critical. In this wise the analysis should take into consideration the characteristics of the data including the type of measurement scale used (e.g., continuous, binary, or categorical), the experimental units, and how the data were collected.

Depending on the type of research design used for data collection, different types of data can be obtained. When the data for the study can be categorized into a number of different groups, referred to as clusters or groups, then the obtained data are structured as clustered or hierarchical data. Each cluster contains multiple observations, giving the data a nested or hierarchical structure, with individual observations nested within the cluster. The key feature of clustered or hierarchical data is that observations within a cluster are "more alike" than observations from different clusters. Clustered or hierarchical data are often obtained in behavioural science researches, but hardly is its analysis discussed explicitly in studies on child discipline.

The observations within a cluster are more alike than observations from different clusters. This fact leads to correlation between the observations within the same cluster, which is referred to as intracluster correlation. In other words, observations within a cluster are correlated, whereas observations from separate clusters are regarded as independent. Since observations within a cluster do not contribute completely independent information, the "effective" sample size is less than the total number of observations from all clusters.

In nested data structures, observations are organized in non-overlapping groups, and arise naturally from numerous data collection schemes. These structures occur when individuals are observed over time (longitudinal repeated measures data) or when a field or area is subdivided into smaller plots on which a treatment is applied (split plots); or when a stratified sampling scheme is used, such as when sampling students within schools within districts (multilevel data).

When data are organized in this manner the observations are no longer independent, so any statistical model used must allow for a more general dependence structure where observations belonging to the same group can be correlated. Hierarchical linear models (HLMs), also referred to as multilevel models, mixed effects models, random coefficients models, and random effects models, allow for such a dependence structure.

HLMs incorporate parameters associated with the global trend (the fixed effects) and parameters associated with the individual observations (the random effects) that govern the variance-covariance structure of the model. Compared to the linear model, additional complexities are introduced in the process of both model fitting and model checking due to the dependence structure and the incorporation of explanatory variables from each level of the data hierarchy.

Until now, statistical methods applied to handle clustered data have not been well developed or widely understood compared to statistical methods for independent data. Therefore, in many studies that generate clustered data the simplest approach have been adopted, namely to ignore the clustering and treat the data as if all observations were independent. Researchers at the time did not always consider the implications of the assumptions that they made about moving variables from one level to another as they combined data about individuals to the group level and analyzed the data based on the number of groups. This approach is inconsistent, because it eliminated the variability of individuals within their groups from the analysis.

Another approach also considered was to disaggregate variables conceptualized at a higher level, for example the size of an organization, and include them in an analysis conducted at the lower micro level. The problem with this approach is, it treats the characteristics of organizations as if they were individuals in the study and produce different results when the data is analyzed separately at the micro or macro level. Not considering the successive nesting of individuals within groups can lead to underestimation of model parameters, and lead to erroneous conclusions (Raudenbush & Bryk, 2002).

2.2 Earlier applications of hierarchical models

Earlier methodological work on hierarchical models includes the works of Wong and Mason (1983), who examined how differences in national economic development interact with adult educational attainment to influence fertility rates. Their research combined economic indicators collected at the national level with household information on education and fertility. Both households and countries are units in the research with households nested within countries and the basic data structure being hierarchical (Anderson et al., 2002; Goldstein, 1991). Using data from fifteen World Fertility Survey (WFS) countries, Entwistle et al., (1986), studied contraceptive behaviour of couples as a function of socioeconomic origins at the individual level, of the gross national product per capita (GNP), and of the family planning effort at the country level. Crane (1991), tested the epidemic theory of ghetto and neighborhood effects on dropping out and teenage childbearing, drawing data from the 1979 Public Use Microdata Sample (PUMS). Although Crane's analysis does not fall within formal hierarchical modelling, his insights into the functional relationship between neighborhood quality and social problems are valuable to multilevel modelers working on the same topic.

Recent years have seen an increased number of applications of hierarchical models for binary data. For instance, Rountree and Land (1996), applied hierarchical modelling and reported distinctive differences between a general perceived risk of crime and a burglary-specific fear. They based their analysis on a victimization survey collected in Seattle, Washington, in 1990. In the dataset, more than 5000 individuals were clustered into about 300 city-blocks, which were in turn clustered into about 100 major blocks.

2.2.1 Data reduction to independent observations

Earlier approaches to dealing with hierarchical models included reducing data to independent observations and using fixed effects regression models. These methods however, did not explicitly accounted for clustering as methods for handling clustered data have not been well developed or widely understood compared to methods for independent data. Several studies that generate clustered data used the simplest approach of ignoring the clustering and treated the data as if all observations were independent. In these studies, the data obtained from individuals were pooled to create a single large dataset.

Another widely used approach to addressing the above phenomenon (*reducing data to independent observations*) consisted of two stages. The first stage involved reducing the multiple observations in a cluster to a single observation by taking a suitable summary (i.e., taking the mean of all the observations in each cluster). The resulting data points, all from different clusters, were thus regarded as independent. The reduced data was then analyzed using standard methods for independent observations (such as a *t* test). This approach was used in a number of studies involving optical recordings by Virmani et al., 2006 and Willeumier et al., 2006 in all aspects of their analyses. These studies involved examining individual synapses, averaging the data from each experiment and reducing multiple observations to a single observation for each experiment. By this method, the correlation associated with data clustering was removed and no longer needed to be accounted for. Analysis that reduced data to cluster means need only to assume that separate experiments are independent.

This approach, however, has limitations especially if there are unequal numbers of observations per cluster. In such a situation an unweighted method of second stage analysis may not be the most appropriate. “Simply taking the mean of each cluster and then comparing these values by a *t* test, for example, does not automatically take the unequal number of observations per cluster into account. Clusters with more observations could be expected to contribute more information and, thus, should be given more weight in the analysis” (Willeumier et al., 2006).

Another limitation is the loss of specific individual information as you take the mean of observations in each cluster. Thus vital specific information about individuals are disregarded by reducing a cluster to its mean value as a tradeoff to eliminate clustering. This loss of information results in weakening the power of the analysis based on cluster means compared to an analysis that includes information on the individual observations.

2.2.2 Fixed effects regression / ANOVA approaches

These approaches include the cluster effect as a factor in a standard regression model. In datasets where only one of the groups being compared is represented in each cluster, a fixed effects regression would not be used as there would be no within-cluster comparison of the groups and hence, insufficient information to estimate both the group effect and a fixed effect for each cluster. This is true for all such datasets, irrespective of whether the data are normally distributed, skewed, binomial, etc. For a normally distributed data the linear mixed model would be the most effective for analysis. On the other hand, in studies where each cluster contains observations from both groups being compared, a fixed effects regression approach may be suitable. In this case a within-cluster comparison of the two groups can be made.

If a model consist of both the group effect and fixed effect for each cluster, the analysis basically would involve controlling for the cluster effect and estimating the group effect at a fixed level of the cluster (ie within each cluster). With this approach, one has to determine the appropriateness of applying a fixed effect or random effect for clustering. Applying a fixed effect would mean the results of the analysis are strictly applicable only to the particular set of clusters in the study. On the other hand applying a random cluster effect, would mean the clusters are regarded as a random sample from a wider population of clusters, and hence the results can be generalized to the wider population. This approach also allows for the inclusion of an interaction effect even if some clusters contain observations from only one group. Furthermore, if the interest is in the group comparison within a given cluster, then the random effects approach allows for the incorporation of information from all of the clusters whereas the fixed effects approach make use of information from only one cluster.

For datasets where clusters contain observations for groups, applying a fixed effect approach to address issues of clustering provides lower standard errors, especially when you have a small number of clusters. When the data are normal distributed or can be transformed to normality, then a normal regression (ANOVA) approach with a fixed effect for clustering and an effect for group can be applied. If the data is not normally distributed a generalized linear model could be used.

2.3 Methods that account for clustering

Methods that take clustering into consideration can be classified into two broad groups; methods that adjust existing tests to account for clustering and methods that use modelling approach to account for clustering.

2.3.1 Methods that adjust existing tests to account for clustering

Methods that adjust existing tests to consider clustering are applied in hypothesis testing of no difference between two groups of observations (ie null hypothesis). Applying these tests depend on the distribution of the data. "If the data is a normally distributed clustered data, the method adjusts the standard two-sample *t* test by an additional factor designed to take account of the intracluster correlation. In this case there is a modification of the *t* test to account for data clustering. If the data is a binary response data, a similar approach involving the adjustment of the usual X^2 statistic would be made" (Donner & Banting, 1988). Other methods, (rank based tests), have been proposed to handle non normally distributed data by Rosner et al. (2003) and . Rosner and Grove (1999). For datasets where a cluster has observations from two groups for comparison, rank-sum tests have been developed (Larocque et al., 2010; Rosner et al., 2006) and where the observations are paired and clustered, signed-rank tests have also been developed (Datta & Satten, 2008; Rosner et al., 2006).

2.3.2 Methods that use modelling approaches and considers clustering

When other covariates (e.g. age of household head, wealth index of household and ethnicity of head) are to be considered in the analysis, modelling approaches that take clustering into account are generally more useful. When there is clustering in data the assumption of independent observation required for ordinary logistic regression to be valid becomes violated. Two approaches normally used to address this problem are using generalized estimating equations (GEEs) and mixed effects models (MEMs). The GEEs deal with clustering by allowing observation within the cluster data to be correlated by specifying a working correlation structure. The MEM approach, on the other hand, accounts for clustering through the inclusion of a random cluster effect in the analysis model (Stirateli et al., 1984).

If one's interest is to measure the average effect of the covariate on the outcome then it is best to use the GEE. On the other if the interest is to measure cluster-varying covariates then the MEM approach is the most appropriate choice. The choice to use either one is therefore influenced by whether the covariate of interest is fixed or varying within the clusters.

Another important factor to consider in modelling clustered data is using the appropriate link function to transform the dependent outcome so that it can be modeled as a linear function of a set of predictors (Azen & Walker, 2011; Hox, 2010). If the link function being applied is identified as identity link then it does not matter which of the two approaches above to use for parameter estimation as the estimated parameters will coincide in this instance. However, if the link function is a logit one then the appropriate model to use is the mixed effect model. In such a case, a specific link function is used to link the probability distribution of the random component (outcome) and the explanatory categorical outcome model. The link function, links the expected value of the random component of Y (E(Y)) to the linear model.

"The GLM approach to dealing with outcomes that result from sampling distributions other than the normal distribution is to incorporate the necessary transformation of the dependent variable and choice of appropriate sampling distribution directly into the statistical model. In other words, if we choose a linear model for a binary outcome (which is assumed to be sampled from a binomial distribution) without transformation, the set of predictors would be unlikely to provide good predicted values of the outcome (Y)" (Hox, 2010).

"The link function therefore transforms the outcome Y in some appropriate manner, depending on its sampling distribution, so that its expected value can be predicted as a linear function of a set of X predictors" (Azen & Walker, 2011). For continuous outcomes, the transformation of Y is not required as it is assumed to be sampled from a normal distribution with a corresponding identity link function that results in the same expected value for Y. The GLM therefore can be used to investigate non-continuous, categorical outcomes through the use of link functions and alternative forms of sampling distributions.

This study focused on mixed effect models because one of the interests was to find if the effects of the covariates or predictors of interest are fixed or varying within clusters.

2.4 Multilevel modelling of hierarchical data

Multilevel models enable examination of relationships at multiple levels of a data hierarchy. Such datasets with hierarchies or higher levels are identified from single-level datasets by the nesting of individual observation within higher level groups or within individuals of the dataset consisting of repeated measures. For example, with single-level datasets, participants are typically selected through simple random sampling with the assumption that they have equal chances of inclusion in the dataset and do not belong to any higher order social groups that might influence their responses. With multilevel data analysis, however, the grouping of participants as a result of the sampling scheme being used, (e.g. neighborhoods selected first and then individuals selected within neighborhood), is the focus of the conceptual model and theory of multilevel modelling (Kreft & De Leeuw, 1998).

2.5 Assumptions for fitting multilevel models

Multilevel models have similar assumptions as other major general linear models like ANOVA. These general linear model assumptions include the assumption of normality that states that the error at every level of the model are normally distributed and the assumption of homoscedasticity, also known as homogeneity of variance, that assumes equality of population variances. Furthermore, the assumption of independence states that cases are random samples from the population and that scores on the dependent variable are independent of each other.

However, for multilevel models, because of the hierarchical nature of their design, some of these assumptions are modified. (Salkind & Green, 2004) so they can be extended to nonlinear relationships as is the case in our study.

Multilevel modelling, however, deal with cases where the assumption of independence is violated and assumes that the residuals at the different levels are uncorrelated likewise the errors at the highest level. To fit a multilevel model then, one first has to determine the distribution for the random cluster effects, as a misspecification of the distribution could have a substantial impact on the results (Litiere et al., 2007).

Multilevel models are subject to less stringent missing values assumptions and they estimate the cluster-specific or the effect of covariates on the outcome conditional on the value of the random effect. These models are therefore suited to estimating the effect of covariates which remain constant within clusters, (Neuhaus et al., 1991).

2.6 Multilevel models with discrete outcomes

The type or nature of outcome or dependent variable of a clustered data is of great importance in estimating accurate parameters. The outcome of interest in most survey data may not only reflect continuous measure, but also discrete or categorical in nature. Multilevel modelling, using continuous and categorical outcomes have been studied by many authors who have identified a number of important conceptual and mathematical differences between models for continuous and categorical outcomes.

Models with categorical outcomes require different underlying mathematical models and estimation methods for analysis. Because of these differences, they are often more challenging to investigate compared to models with continuous outcomes. With respect to multilevel modelling, models with categorical outcomes use different estimation procedures, which usually take longer time to converge on a solution compared to typical continuous outcome models. Social science data often contain variables which are not continuous. Most datasets are dichotomous e.g. deciding whether or not to vote or dropping out a group or membership of different groups (eg religious affiliation, race/ethnicity).

Fewer studies have examined the accuracy of estimates, sample size or power analysis in discrete dependent outcomes in multilevel modelling. Categorical or discrete outcome models fall in the class of generalized linear models - dichotomous and polythomous. They have between two or more discrete alternatives. Multilevel modelling with discrete response variables, statistically relate each discrete response made by the individual to the attributes of the individuals and the attributes, of alternatives available to the individual and the context in which the individual lives. Such a model estimates the probability that an individual choose a particular alternative and also predict how people's choices will change under changes in demographic and or attributes of the alternatives. Also known as categorical dependent outcome models, they are classified according to the number of available outcomes or alternatives.

Prominent types are binomial choice models (dichotomous) with two available alternatives or outcome and multinomial choice models (polytomous) with three or more alternatives or outcomes. They can further be classified as models with standard logit, which assumes no correlation in unobserved factors over alternatives and models that allow correlation in unobserved factors among alternatives or outcomes. They have the assumptions of normal distribution and homoscedastic errors are violated (Hox, 2010; O'Connell et al., 2008).

Other common features of discrete choice models include having a choice set and defining choice probabilities. The choice set must meet three requirements: (1) The set of alternatives must be collectively exhaustive ie it includes all possible alternatives, (2) The alternatives must be mutually exclusive meaning choosing one alternative means not choosing any other alternative and finally, (3) the set must contain a finite number of alternatives. That is, the dependent variables have a finite number of values. Defining choice probabilities means for example the multinomial model specifies the probability of choosing a particular alternative or category with probability expressed as a function of observed variables that relate to the alternatives.

This study focused on the multilevel logit models where we considered estimation of cluster-specific effects in two modelling approaches to model: (1)

dependent binary and (2) multinomial outcomes with hierarchical structures. The two modelling approaches are described in the next chapter explaining the mathematical models behind them. The performances of the models were examined through simulation studies and then applied to model the MICS 4 dataset on child discipline practices in Ghana.

2.7 Model evaluation through Simulation

Simulation studies provide powerful conclusions for correlated or longitudinal response data particularly for relatively small samples for which asymptotic theory does not apply. Simulation studies are useful tool for examining the robustness of a given statistical test or the performance of statistical methods. There have been a number of studies that compared estimating methods using simulation studies. Feng et al., (1996), used simulations to compared LMM and GEE; Datta and Satten, (2012) also compared their test with both the standard Wilcoxon test ignoring clustering and the Wilcoxon test on cluster means. Rosner et al., (2006), compared their test with LMM and the Wilcoxon signed rank test. Larocque et al., (2010), compared their proposals with test conducted by Rosner et al. (2006) and Datta and Satten (2012). The focus of these studies was to compare the robustness and performance of their developed models using simulated datasets.

2.7.1 Comparing estimation methods for hierarchical data using simulations

In terms of robustness that is, how liberal or conservative a given test is, one first generates datasets according to a model that specifies no difference between the two groups being compared. A null hypothesis of no difference between the two groups at 5% significance level is then tested. Accordingly, we should see the null hypothesis rejected for about 5% of the datasets. A test that rejects the null hypothesis, greater than 5% of the time, is considered too liberal, and one that rejects it less than 5% of the time is considered too conservative. Secondly, the power of a given statistical test, which indicates how effective a test can identify that differences exist between two groups is also investigated. To achieve this, data is generated under an alternative hypothesis, specifying a real difference between the two groups. Given two tests that maintain the correct 5% size under the null hypothesis, we would prefer the one that correctly rejects the null hypothesis for a higher proportion of datasets generated under the alternative hypothesis, since it has higher power.

2.7.2 Using simulations to assess the performance of statistical methods

In terms of assessing the performance of a statistical method, one first generates outcomes or covariates based on a chosen simulation model. The simulated datasets are then analyzed using statistical methods of interest and the results are compared to the true parameters specified in the simulation model.

2.7.3 Simulating hierarchical dataset with categorical outcomes

For hierarchical modelling, it is important to have appropriate methods for simulating correlated binary or multinomial data along with associated predictors. In the absence of clustering simulating categorical outcomes is relatively simple. This involves determining the probability of experiencing the event of interest for each subject based on their covariate values and according to the chosen simulation model.

Simulating binary outcomes in a clustered data setting is more challenging and two broad approaches are usually considered. Firstly, the data is generated under a marginal simulation model by specifying the marginal or population averaged parameters. Secondly, a condition simulation model is used, where the condition or cluster-specific parameter are specified.

In the marginal approach the marginal success probabilities are allowed to depend on item level covariates and hence vary widely within a cluster. This restriction greatly limit the range of correlations that can be accommodated (Oman & Zucker, 2001). The conditional simulation is used to evaluate the uncertainty in the numerical model parameter values and the translated uncertainty in the model response.

2.8 Studies on Child discipline

"Child discipline is the guidance of children's moral, emotional and physical development, enabling children to take responsibility for themselves when they are older" (Holden, 2002; Wissow, 2002). Child discipline has to do with educating children on the values and actions of what is acceptable and what is not acceptable in their family and in society. When children are praised for exhibiting what is considered as good behaviour or not exhibiting what is considered as bad behaviour, it is termed as positive discipline. Positive discipline is also the situation where for example a child is smacked for demonstrating what is considered as inappropriate behaviour. The objective of Positive discipline is to assist the child to know why certain conduct is unacceptable and other conduct is acceptable. On the other hand, negative discipline aims at enforcing compliance to avoid something unpleasant.

"The discipline and punishment of children by parents is among the most commonly investigated topics in developmental psychology. Studies on socialization show that the place of early child discipline is very critical especially the various processes used to teach children the values to become competent adults. Holden

(2015b), concluded that "the types of disciplinary techniques used by parents reflect a core ingredient of those parents' approach to child rearing and have relationship with children's outcomes".

"Research shows that children who experience corporal punishment and psychological aggression may exhibit a range of behavioural, emotional and psychological problems, such as antisocial behaviour, violence, anxiety, posttraumatic stress disorder, and anger" (English et al., 2008; Gershoff, 2002; Straus, 2001).

Studies on child discipline dates back to the early 1960s and 1970s when American evangelist, Christian James Dobson, advocated for the return of a more conservative society by promoting biblical parenting. In his book titled "Dare to Discipline", published in 1977, he advocated for spanking of children up to the age of eight (Dobson, 1977). Subsequent theories on child discipline have been developed by researchers in Europe and North American countries using predominantly data from white working class and middle-class families (Dodge, 2002). These studies provided the first detailed data on the types of discipline methods used by parents, why they use them, and how discipline is associated with children's behaviour (Gershoff et al., 2015).

These studies, however, did not expand the population of interest beyond families in Europe and North American countries and as a result led to criticisms that the theories and recommendations developed are not universal (as often assumed to be) and may not apply to other countries and cultures (Dasen & Mishra, 2000; Rogoff, 2003). These concerns led to interest into research in other cultures from non-western and non-industrialized societies. Broadening the research beyond Europe and North America to include families from other cultural, religious and economic backgrounds revealed important variations between Asian and European families based on beliefs and cultural practices (Erbacher, 2002; Solis-Camara & Fox, 1995).

Research, further showed that the goals, values and types of discipline strategies used by parents vary based on a number of factors including ethnic background (Garcia et al., 1995). While African American parents place emphasis on interdependence, security, collective goals and common interests and perseverance in the context of adversity (Hill & Bush, 2001), European American parents, especially middle-class parents, tended to approve individualism, individual achievement, competition and material well-being to a greater extent than others (LaReau, 2003).

“Endorsement of these values is associated with greater use of democratic, parenting strategies that emphasize reasoning and discussion” (LaReau, 2003). “For both African American and European American families, the predominant parenting practices endorsed within each ethnic group has been associated with positive and adaptive developmental outcomes, including higher levels of academic achievement and lower levels of mental health problems” (Hill & Bush, 2001).

“Historical research on child discipline suggests that there has always been a great deal of individual variation in methods of discipline” (Pollock & Linda, 1983). Orme (2001), argued that “Children in medieval times were treated differently from adults in legal matters, and authorities at the time were as concerned about violence to children as they were to adults”. In his article, "Childhood in Medieval England," he states, "Corporal punishment was in use throughout society and in homes, although social commentators criticized parents for indulgence towards children rather than for harsh discipline”. Corporal punishment was the norm at the time as the medieval world was considered a dangerous place, and harsh measures were required to prepare a child to live in it. Biblical views on child discipline mention the importance of disciplining children, as opposed to leaving them neglected or unruly, in several verses. The most often paraphrased is Proverbs 13:24, "He that spareth his rod hateth his son: but he that loveth him chasteneth him betimes".

David Robinson, a writer for the *Colonial Williamsburg Journal*, noted that "during colonial times in the United States, Puritans permitted their young children to play freely and older children were expected to swiftly adopt adult chores and accountabilities, to meet the strict necessities of daily life" (Fleming, 2012).

John Locke, an English physician and philosopher, in his 1690 essay on *Human understanding*, argued that "children can be compared to a blank tablet (tabula rasa) at birth and not inherently full of sin". In his second article titled *Some Thoughts Concerning Education*, Locke argued that "the task of the parent was to build in the child a strong body and habits of mind that would allow the capacity of reason to develop. The parents could reward good behaviour with their admiration and punish bad behaviour with disgrace i.e withdrawal of parental approval and affection as opposed to beatings" (John, 1693).

Experts in child rearing, in the early twentieth century, took a different approach on child development and promoted appropriate ways for child discipline. John Watson, a behavioural psychologist in 1924 argued that "parents should train their young children by rewarding good behaviour, punishing bad ones, and following precise schedules for food, sleep, and other bodily functions". Before then, a 1914 U.S. Children's Bureau pamphlet, *Infant Care*, recommended a strict schedule and cautioning parents not to play with their babies.

2.8.1 Theoretical contributions to child discipline methods

There has been a number of theoretical contributions to the study of child discipline methods. These contributions have traditionally come from research on socialization of the parent-child relationship and are obtained from theories that emphasized the importance of relatively specific and single social processes (Maccoby, 1992). There are also theories that include reinforcement (Patterson, 1982), social cognition (Bugental, Johnson, New, & Sylvester, 1998) and emotion regulation (Cole et al., 1994). Other theories emphasize on the dynamic interplay of the various social processes on parent child relationships (Cummings et al., 2000). Additionally, there are theories and worldviews that buttress research on the influence of culture, ethnicity, socio-economic status and community characteristics on parenting, family dynamics and children's development.

2.8.2 Socialization of the parent-child relationship

According to Grusec and Goodnow (1994), "if the goal of discipline is to promote children's compliance and internalization of parental and societal values, then an appropriate combination of power-assertive and inductive disciplinary techniques should be applied to successful internalization". Hoffman (2002) observed that "The theory of moral internalization attempts to address how societal norms and parental values (initially motivated by external forces like, fear of sanction) eventually come to acquire an internal motivational force". Hoffman, argued that "disciplinary encounters with parents are central to this process of moral internalization and what happens in a disciplinary encounter is likely to influence whether or not children internalize norms and subsequently behave in a way that is consistent with these norms".

Hoffman (1983) further proposed that “the use of physical force, deprivation of privileges or possessions, direct commands or threats are detrimental to socialization”. He argued that “these actions do not only produce fear and anxiety in the child but also provide a model of aggression to the child”. Again, these “heightens the child’s view that the moral standard is external to the self, and direct the attention to the consequences of the behaviour for the self rather than for other people” he concluded. Hoffman (1983), again argued that “such oriented inductions promote internalization because they develop the child’s empathic capacities and induce negative feelings from which the child cannot escape even when the parent is no longer present. These, inductions are unlikely to produce high levels of anxiety or fear, and so the child is more available to attend to the process the information embedded in the parent’s inductive statement”.

Kochanska and Thompson (1997), concluded from their studies that “power-oriented and forceful discipline creates anxiety in children and interferes with the effective processing of parental message about behavioural standards” and thus undermines the process of internalization.

Further studies show that when children are over protected, especially by the mothers, it increases their anxiety (Van et al., 2009; Verhoeven et al., 2011). In their study Robinson and Cartwright (2008), observed that the application of severe and ineffective disciplinary methods by mothers increase the anxiety of the children. Sadoughi et al. (2007), also found that “the susceptibility of sensory processing, in interaction with inappropriate parenting methods lead the individual to levels of anxiety”.

2.8.3 Reinforcement and emotion regulation

Theories on reinforcement and emotion stem from learning and social learning theory and considers the strategies of reinforcement and punishment as central to learning and socialization. The learning theory suggests that "consequences of behaviour strengthens or weakens behaviour in the future". That is, behaviours that are rewarded are maintained by the child, while behaviours that are punished drop out (Eisenberg & Valiente, 2002). Within this context, Domjan (2000), defined punishment as "the presentation of an aversive stimulus or the removal of a positive stimulus". He argued that "For punishment to be effective for behavioural change, it should be administered immediately after the transgression". That is, administering punishment following a specific behaviour is likely to decrease the possibility of that behaviour recurring. Holden (2002), argued that for physical punishment to successfully suppressed a behaviour, it should be severe enough to serve as a negative consequence for the child.

Modern social learning theory (Eisenberg & Valiente, 2002), hinges on the fact that "children learn through observation and imitation of models in their environment". The actions or inactions of parents serve as important source of information to children about behaviour expectations and possible consequences for various behaviours. Again, Social learning theorists argue that the use of physical punishment on children, models aggressive behaviour for children (Straus, 1991). According to Smith et al. (2005), "Parental discipline is nested within a wider system of relationship in the family, as well as overarching systems of social and economic influences, which impede or facilitate effective parenting". Ethnicity and culture have been observed as important factors that influence the application of different discipline methods. Brody and Flor (1998), reported that "African-American parents use high levels of control and physical punishment along with high levels of affection and acceptance, while European-American parents use democratic parenting practices more often".

In their studies, Deater-Deckard et al. (2003) and Dodge (2002), observed significant positive correlation between harsh punishment and acting-out behaviour for African-American families, but not European-American ones, in the USA. Lerner et al. (2002), suggested that "'because the meaning of parents' behaviour is critical in determining its effect on the child and because community standards determine the meaning of parents' behaviour, physical punishment is likely to have very different consequences in different communities'".

"Apart from ethnicity and cultural background, socio- economic status also impacts parenting beliefs and practices. Social economic status defined as comprising family income, parental education levels, prestige of parents' occupations, wealth and material possessions, influence parents' theories about child development, the characteristics parents wish to develop in their children and their beliefs about parenting" (Holden, 2002; Hoffman et al., 2002). Social economic status has great impact on parenting beliefs and practices due to its association with families' access to material resources (Duncan & Maguesson, 2002).

Other key factors that have emerged to influence parenting practices include the social capital and knowledge that parents bring to parenting (Hoff-Ginsberg & Tardiff, 1995) and between parents' occupational conditions on their beliefs about important values to instill in their children, influence parenting practices (Hoffman, 2002).

"Consistently, parents of families with lower levels of socio-economic status more often endorse and use physical discipline, hierarchical and authoritarian relationship styles and focus on obedience, conformity and maintaining order" (Hoff-Ginsberg & Tardiff, 1995; Pinderhughes et al., 2000). "In contrast, parents of families with higher levels of socio-economic status more often endorse and use psychological punishments, guilt induction, egalitarian relationships between parents and children and focus on developing independence and questioning authority"(Hoff-Ginsberg & Tardiff, 1995; LaReau, 2003).

Another important characteristic of parental discipline is the child's awareness and recognition of the parent's disciplinary message (Holden, 2002). Characteristics like gender, temperament, age, and history of interaction between child and parent, are known to impart on Children's reactions to the type of disciplinary strategies adopted by parents. (Holden, 2002).

2.8.4 Parenting styles and Child discipline

Parenting styles play important role in the overall emotional climate in the home. They are generally considered a helpful gauge of a child's success, the level of control exercised in parent-child relationship and communication methods. The type of parenting style adopted by parents in the upbringing of their children also provide valuable information in understanding effective child discipline (Maccoby, 1992).

Studies on parenting styles have come out with two key dimensions: parents who are high or low in control or demand and parents who are high or low in warmth or responsiveness. Depending on where parents lie along these dimensions, Halpenny, Nixion, and Watson (2010), categorized them as authoritative (high-control, high-responsive) or authoritarian (high-control, low-responsive) or permissive-indulgent (low-control, high-responsive) or permissive-neglectful (low-control, low-responsive). "The effect of the intended discipline may vary depending on whether it is in the context of a warm and responsive parent-child relationship" they concluded.

Parents' or caregivers who apply hash punitive discipline methods can be likened to authoritarian parents who exercise firm control, but in a rejecting manner. Such parents or caregivers do not discuss the misbehaviour of children, but are more likely to mete out punishment. Findings from research show that when children are raised by parents with authoritarian behaviour such children tend to be more hostile, aggressive and less popular with their peers. Such children also have difficulty of being independent, have less academic success and engage in substance use as teenagers (Lamborn et al., 1991; Steinberg et al., 1992).

Child discipline methods which are violent in nature can either be physical or psychological or a combination of both (Turner & Muller, 2004). Violent child discipline methods also called corporal punishment involves the application of physical force such as spanking to force children compile. It is also identified as discipline where physical force is applied with the intention to cause pain or discomfort. Application of violent physical discipline methods, have been associated with many adult mental health problems, delinquency and adult criminal behaviour. Similarly, psychological aggression discipline methods which comprises the application of guilt, humiliation, withdrawal of love or emotional influence to control children have also been proven to have consequences for children. Lau (2010), found that Chinese fathers who applied physical control on their sons, had more physically aggressive sons and mothers who used psychological aggression methods to control their daughters ended up with more physically and emotionally aggressive daughters.

Non-violent child disciplinary practices involves taking away privileges or explaining why exhibited behaviour is wrong. It also include ignoring and use of distraction before the wrong behaviour escalates. Authoritative parents tend to apply these methods to monitor their children closely. They set clear standards and high expectations for the children and allows communication between parent and child.

Authoritative parents tend to be understanding and supportive. They set boundaries and institute appropriate consequences if the child does not behave. Children raised by authoritative parents tend to experience greater academic success, are less hostile, have higher self-esteem, and show more purpose and independence.

2.8.5 Effects of parenting styles on children

Research has shown that inappropriate parenting styles result in various consequences for children and parents. Sayyed et al. (2008), found that children who tend to be aggressive and not law abiding usually come from homes where permissive parenting styles are used. Rahmani et al. (2005), also found that students with parents of non-authoritative parenting style had different behavioural problems. The destructive effects of authoritarian and permissive parenting styles and positive effects of authoritative style on children was also observed in the study of Khouynejad et al. (2006) where they found that authoritarian and permissive parenting styles led to a feeling of loneliness in girls. Sadrosadat et al. (2005) also showed that families with authoritarian and permissive parenting styles have low familial function. Bagherpour et al. (2007), showed that mental health and educational achievement is more prevalent among the children nurtured by authoritative style comparing to children nurtured by authoritarian or permissive styles.

2.8.6 Observed inconsistencies in parenting strategies in literature

There have been some inconsistencies in the literature about the parenting strategies often used by families. For example some researchers are of the view that Latino families tend to allow permissiveness, nurturance and egalitarianism (Delgado, 2002; Durrett et al., 1975; Vega, 1980), while others suggest that they are more likely to use authoritarian parenting practices (Chilman, 1993; Harrison et al., 1990). There are still other researchers who think they are equally likely to support democratic and authoritarian parenting strategies but not permissive (Martinez, 1988). These observed inconsistences may not be due to differences in the socioeconomic status of families alone, but are probably due to other factors like differences in research design, acculturation and confounding ethnicity with social economic status and community context (Nancy, 2006).

2.8.7 Theoretical and conceptual frameworks for parental discipline

The study of parental disciplinary practices can be situated within diverse theoretical and conceptual frameworks such as learning and social learning theory, the theory of moral internalization, parenting styles approach, and ecological approaches.

2.8.7.1 Learning and social learning theory

The central idea here is reinforcement of what is accepted as good behaviour and punishment for what is seen as wrong behaviour. The learning theory proposes that the end results of a behaviour either strengthens or weakens the demonstration of the same behaviour in the future. "Behaviours that are rewarded continue in the child's repertoire, while behaviours that are punished drop out" (Eisenberg & Valiente, 2002).

2.8.7.2 Theory of moral internalization

Hoffman (2002), came out with a theory that "moral internalization addresses how societal norms and parental values which are motivated by external forces (e.g. fear of sanction), eventually come to acquire an internal motivational force". He argued that "disciplinary encounters with parents are central to this process of moral internalization and what happens in a disciplinary encounter is likely to influence whether or not children internalize norms and subsequently behave in a way that is consistent with these norms".

2.8.7.3 Parenting styles approach

In this approach, illustrated by Baumrind (1991) and Maccoby (1992), two key dimensions of parenting styles are identified: Parents who are high or low in control or demands, and parents who are high or low in warmth or responsiveness. Based on where parents lie along these dimensions, they are categorised as: authoritative (high-control, high-responsive); authoritarian (high-control, low-responsive); permissive-indulgent (low-control, high-responsive) and permissive-neglectful (low-control, low-responsive).

2.8.7.4 Ecological /systemic approach

In this approach, the consequence of discipline varies depending on whether discipline was meted out in the context of warm and responsive relationship between parent and child or otherwise. Parental discipline is known to be influenced by a multiplicity of factors which either facilitate positive parenting of worsen it (Smith et al., 2005). Factors like culture and ethnicity have been identified as important influences on the results of different discipline practices. Another major aspect of the ecological and systemic approach is child effects on parenting. In explaining this process, Patterson (1982), developed a model of coercive family processes, which identified that there are "bi-directional parent–child interactions that contribute to the development of aggressive behaviour in children".

This research, based on the ecological and systemic approach, holds the idea that a number of factors within an individual (parent and child), within the family setting and beyond, influences parental discipline.

2.8.8 Discussion of Bronfenbrenner's Ecological Conceptual Framework

The Bronfenbrenner's ecological conceptual framework postulates that "parental practices are developed within interactive effects of multiple and interdependent systems". In his studies, Bronfenbrenner (1997), situated Ghanaian discipline practices within a context of four ecological systems: the micro-system, the meso-system, the exo-system, and the macro-system.

The micro-system, is the immediate surroundings that affect the delivery of child discipline. While it is easier to identify the effects of cultural beliefs on parenting, when children interact with immediate families, it becomes more complex when the perception about child discipline interacts with the child's immediate external systems. Parental discipline is influenced by other systems just like children interaction with outside entities, (day care or pre-school or face to face contact), which influence them. In other words, what is practiced at home can be contested by new outside influences. Studies show that, the immediate home surroundings have a lot of influence on how to train the child. But as other systems interact, they provide other ways as to how children should be raised. For example, as parents interact with daycare or health centers they are exposed to new ideas and suggestions which they can integrated with their principles of parenting their children.

The meso-system, refers to "the interaction between parental values/beliefs and the various family levels that the child interacts with". For example, parent-teacher collaborations at preschool can influence parent-child interactions at home. These include, "the interrelationships, the initiatives of the child, and the parents involvement in linking the home, other family members and the school". These factors "play important roles in determining the quality of the child's meso-system".

The context within which parental disciplinary choices such as values, beliefs, and ideologies are made, is the cultural context. This, however changes, as the child grows older and comes into contact with other systems outside of the home. As interactions with the other systems go on, it begins to have influence on parental choice of disciplinary practices. Hence, where physical forms of discipline used to be applied, parents begin to integrate other forms of discipline preferred by the immediate family context. The impact of these other effects could be a source of support or stress to the family system in raising the child.

The third context is the exo-system made up of social structures that indirectly impact the micro-systems of parents. In this context, the relationship among settings (neighborhood, workplace, parent's social network) is influenced by forces beyond the child. When parents interact, they learn of the societal disciplinary practices in their new environment that may be different from what they know, and may be influenced to amend their practices even though they have their own values and beliefs.

The macro-system, is a combination of the three previous systems, including the cultural and religious beliefs of the society/community which are mirrored in the community's political, economic, and educational systems. In this setting the general ideological and organizational patterns reflect the ecology of human development. It represents how broader ideologies interface with how parents discipline their children. For example, this study revealed significant associations between parents' level of education, age-group, and wealth index of household and child discipline practices.

In conclusion, the ecological concept theorizes that different ideologies from religious, societal, education, laws and political systems interface with the values and beliefs that parents have.

2.9 Analysis of Child discipline data

Child discipline studies come under behavioural and social research where interaction between the individual and the social group they belong play an important role. In this context, the individual and the social group they belong to are considered as a hierarchical system of individuals nested within groups and are defined at separate levels of this hierarchical system. "Despite this obvious hierarchical structure, often times its overlooked in analysis leading to errors in interpretation of results of statistical significance tests" (Goldstein, 2003b). That is, most analysis of child discipline data have often not taken into consideration how to relate the characteristics of individuals and properties of groups and structures in which the individual functions. Most studies have applied either descriptive or regression models without considering contextual effects. Stoolmiller et al,. (2004), presented a multilevel modelling approach in their study of parental discipline and child antisocial behaviour. The analysis however, was done in the context that the data was independent ie without clustering or no dependence between the observations collected on the units belonging to the same cluster. The assumption that the data is independence cannot be realistic in this instance. This is because a number of theories highlight the importance of the complex interplay of factors that shape parental discipline responses such as individual child and parent characteristics and behaviours, and more generally the contextual influences within which parenting occurs. Therefore, statistical analysis of such data, should take into consideration the intra-cluster correlation that results from the interplay these factors. Neglect of this factor in analysis can lead to erroneous conclusions concerning the impact of the diverse sources of influence on the outcome variable.

No previous studies on child discipline methods have used multilevel analysis that took clustering into account to establish relationships between the response variables, the covariates and hierarchies in the collected data.

Until the conduct in Ghana in 2006 and 2011 of the Multiple Indicator Cluster Surveys (MICS 3 and MICS 4 respectively), which examined child discipline and violence in the home environment, limited studies on child discipline in Ghana focused on children in schools. These studies were analyzed descriptively and suggested a high percentage use of physical punishment (corporal punishment) and an increased rise in the use of physical disciplinary methods in Ghanaian homes without accounting for clustering in the analysis (Ghana Statistical Service, 2006, 2011). Studies on "Better discipline for Ghana's children" in 2006, by Branund and Clarke, and in 2014 by

Danvers and Schley both used thematic analysis of their data which do not consider clustering.

Knowing the important role clustering play in hierarchical structured data, and the statistical implications of ignoring it in the analysis (such as reduced standard errors and inflated type 1 errors rates (Heo & Leon, 2005), it is critical that it should be considered if precise parameters are to be estimated.

2.10 Factors influencing parental discipline strategies

Research findings indicate a number of factors interact to influence the way parents discipline their children. Identifying the exact nature of the influence therefore becomes an arduous task. Some identified factors, however, include the characteristics of the child, the characteristics of child's parent / caregiver and the environment in which the child and parent dwell. Xu et al., (2007), concluded from their study on factors that influence parental discipline strategies, that "A parent's overall style of child upbringing, including likelihood of using particular type of discipline method, is partly determined by the set of parenting beliefs, goals, and expectations inherent in his or her culture's model of parent–child relations otherwise called cultural capital".

In addition to predicting whether parents will use physical discipline methods or reasoning and other child-oriented discipline techniques, cultural context is also known to play a key role in this determination. When physical discipline is acknowledged and expected by the overall community, parents tend to feel right in applying physical discipline and the recipient child may see it as normative. This research focused on how parental characteristics, beliefs and contextual factors operate to shape and determine parenting behaviour in terms of child discipline taking into consideration the effects of different levels or groups that households belong.

2.10.1 Parental characteristics

Parents use or approval of physical discipline is influence by a wide range of parental characteristics. With regards to gender of parents, findings are inconclusive. While some studies conclude that there are no gender differences (Hemenway et al., 1994; Murphy-Cowan & Stringer, 1999; Nobes et al., 1999), others conclude that mothers tend to use physical punishment more than fathers (Anderson et al., 2002; Straus & Stewart, 1999). The use of higher levels of physical discipline is also found among younger parents (Dietz, 2000; Durrant, 1999; Giles-Sims et al., 1995). Parenting values and beliefs have also been shown to influence the discipline responses of parents (Pinderhughes et al., 2000), with less-educated parents identified as applying physical punishment the more (Durrant, 1999). However, there are exceptions to these findings as other studies that link the application of physical punishment by parents to higher levels of education and other studies suggesting no effect on the level of education (Dietz, 2000). "The more frustration, irritation or anger a parent feels in response to conflict with a child, the more likely they are to use coercive discipline strategies, including physical punishment" (Ateah & Durrant, 2005; Holden, 2015b).

Findings on studies on intergenerational transmission of parenting practices and attitudes show that "parents who were themselves physically punished as children or adolescents have an increased likelihood of physically punishing their own children" (Bower-Russa et al., 2001).

Garcia et al. (1995), found that “higher levels of acceptability of physical discipline of children in 14 countries in the European Union, were reported among men, older parents and less educated parents”. Ghate et al. (2003), also found that parents “who approve physical discipline were five times more likely to use it than those who disapprove”. “ Parents’ own experiences of discipline during childhood can also shape the particular discipline style they adopt when parenting their children” (Bugental et al., 1998).

2.10.2 Contextual factors

Cultural values, social norms and tradition play major roles in how parents raise their children. A number of studies have shown that, the social class, wealth, culture and income of parents have strong impact on the type of discipline methods to apply in child rearing (Annette & Lareau, 2002). Contextual factors, such as family structure, have been found to be associated with parental use of physical discipline methods (Smith et al., 2005). Parents of larger families are more likely to approve of and use corporal punishment than parents of smaller families (Eamon & Zuehl, 2001). Increased use of physical discipline methods is also associated with marital conflict or violence, relationship stress and parenting stress (Coyl-Shepherd et al., 2002; Wilson et al., 2002; Wissow, 2001).

Nobes et al. (1999), compared the effects of psychosocial and economic stresses and family structure on the chances that a family will apply physical discipline methods on families. They found that psychosocial and economic stresses were more significant in predicting the application of physical discipline methods by families than family structure. They also observed that child maltreatment had a high relationship with poverty, poor mental health and discordant marital relations, than the number of parents in the family. Anderson et al. (2002) and Bower-Russa et al. (2001) also found that parents desire to prevent their child from doing something dangerous is a key factor in attracting the application of physical punishment on the part of parents.

Findings from researches on effects of cultural practices on child discipline is inconclusive and largely contradictory. According to Smith et al. (2005), most of these findings on ethnic differences in attitudes towards use of physical punishment are inconclusive and sometimes contradictory. "Overall, there is not a great deal of evidence to suggest significant differences in prevalence or severity rates for the application of physical discipline methods across different ethnic groups and where differences are reported, the effects are very small" they concluded. "The existence of laws prohibiting physical punishment of children, as well as a lower number of child maltreatment deaths, were significantly associated with lower levels of acceptability of physical punishment of children" they added.

2.11 Empirical evidence of variations in use of discipline methods across countries

Studies on variations and differences in attitudes to the use of physical discipline methods across different countries suggest a strong association between parents' approval and actual practice. In Barbados and Korea, for example, where support for corporal punishment is very strong, rates of severe punishment are very high. In Canada and China, where support is lower, its administration appears to be less severe. In Sweden, where support for physical punishment is very low and laws are explicit about its unacceptability, its use is rare.

Smith et al., (2005), pointed out that "the issues related to ethnicity are confounded with a range of other factors and variables that make it difficult to establish the effect or relative influence of culture or ethnic group status". Several studies have investigated the role of religious beliefs as a determinant of parenting attitudes and have also linked them to behavioral intentions (Ellison et al., 1996; Gershoff, 2002). For example, conservative Protestants have been observed to have a greater frequency of slapping their 3-year-old children, as well as holding more positive attitudes about corporal punishment, than other groups.

In summary, contextual factors, such as family composition and structure, reveal no clear effects on the extent to which parents adopt physical punishment as a discipline strategy. More significant in influencing discipline responses reported by parents are the stress levels in the area of marital conflict and relationships. The influence of cultural and ethnic differences on the use of physical punishment is not very clear and research findings have been largely inconclusive and contradictory in explaining patterns of use related to physical punishment.

2.12 Parenting in Ghana

Families play important role in the development of human competences and character. In Ghana, "the sense of belonging to a family and clan is very strong and establishes rights and obligations for all members, including children" (Twum-Danso, 2012). She confirmed this observation in her study where she concluded that "Children grow up in a closely connected extended family network with strong cultural traditions governing their birth, socialization and upbringing. In many communities particularly in rural areas, members of the wilder extended family are expected to participate in the upbringing of children". For many, family does not mean only the nuclear family, but includes aunts, uncles, cousins, nieces, nephews, and grandparents. Childcare is often provided by extended family when parents work or are away from home, and they sometimes assume as much responsibility for raising the children as the parents.

This relationships, traditionally, act as a kind of a social welfare system that ensures that resources are fairly distributed across the different levels of the family network for the survival of all and strengthening of family ties in the process. There is also informal fostering, where children are sent to live with another relatives. In recent times, however, it's been observed that the extended family network is weakening as a result of poverty and rural to urban migration.

In the Ghanaian cultural context, the families play important roles in creating a supportive and protective environment for the child as he/she grows. Parents generally, are considered as the basic unit for child upbringing across the length and breadth of the country. Majority of children live with at least one biological parent. Providing the child's basic needs is considered the most important responsibility of the parents. For children who do not live with their parents or without parental care, the extended family provides some material and financial assistance. Children without parental care are rarely abandoned. The informal fostering system, however, is sometimes abused and children become vulnerable and exploited.

2.13 Empirical studies on child discipline in Ghana

Limited studies on child discipline in Ghana, have sought to explain the practice and parenting styles in the country. Unfortunately, most of these studies did not go into detail analysis as they dwelt mostly on physical punishment and came out with only descriptive analysis of the situation. A survey by Campaign for Female Education indicated that, of the 2,314 parents, students and graduates interviewed, 94% of parents, 92% of students and 89% of female graduates supported the use of corporal punishment in schools. Also, 64% of teachers agreed that corporal punishment must be tolerated (Reported in GhanaWeb, 18 August, 2011). Again, in a survey carried out by Action Aid in collaboration with Songtaba in 2009, it was found that seven out of eight boys thought that "corporal punishment, such as being caned, having their ears pulled or forced to kneel, weed or dig pits, was necessary, while more than a quarter of the girls interviewed said they would absent themselves from school because of the fear of punishment" (Reported in Modern Ghana, 6 September 2011, www.modernghana.com).

A UNICEF report published in 2010 states that 90% of children aged 2–14 experienced violent discipline in 2005–2006. Seven children in ten experienced physical punishment, while 46% of mothers and caregivers thought that physical punishment was necessary in childrearing. Nearly nine in ten children, also experienced non-violent discipline. One child in ten experienced severe physical punishment and 84% experienced psychological aggression. UNICEF statistics, on violence in the family within the period 2005 to 2006, indicated that children with disabilities were more likely to experience severe physical punishment. The statistics further showed that "15% of disabled children aged 2-9 were hit or slapped on the face, head or ears or hit over and over as hard as possible with an implement, compared with 8% of non-disabled children. Forty-seven per cent of girls and women aged 15-49 thought that a husband is justified in hitting or beating his wife under certain circumstances" (UNICEF, 2009).

A Government report involving 4,164 children found that 81% of children experienced corporal punishment in the home and that at school, caning was the main punishment, experienced by 71% of children (IRAD Report, February 1, 2011).

In a report on MyJoyOnline (2010), titled “Scarred with Whips: The agony of Osu Children’s home inmates”, it was found that teachers and caregivers use coporal punishment including caning, kicking and slapping on children. The Report further indicated that “ in care institutions and schools in Ghana, there is widespread application of corporal punishment as a result of which some children had developed fear and dislike for their caregivers, with many others playing truant to escape corporal punishment at school” (Reported in “Scarred with Whips: The agony of Osu Children’s Home inmates”, MyJoyOnline, 2010).

Using face-to-face interviews, diaries and a questionnaire, Twum-Danso (2010), found a high prevalence of application of physical punishment in private schools in her study on children’s perceptions of physical punishment. Out of the 158 children in her sample who participated in the study, she found that “61.4% experienced some physical punishment at the hands of parents or primary caregivers, with 30.4% experiencing only physical methods of punishment at home. Seven in ten of survey respondents said that school was the place in which they were most likely to be physically punished. Caning was the most common method of physical punishment at home and at school”.

A survey by UNICEF (2014), on disciplinary attitudes, revealed “inconsistencies between what adults and children perceive in theory to be the best ways to discipline children and what is actually happening in practice”. It was also observed that “although corporal punishment is generally widespread, both adult and child survey respondents in general believed ‘positive’ discipline techniques to be the best ways to discipline and ‘negative’ techniques, especially involving physical punishment, to be the worst ways to discipline children”.

In his study to seek the opinion of respondents on the best methods to apply in child discipline, Kyei-Gyamfi (2010), found 58% of the survey respondents (adult and child) indicated ‘positive’ discipline methods as the best methods to discipline children. Such ‘positive’ methods included; making sure the child receiving the discipline knows what he/she did was wrong, rules should be explained well to them, and good examples should be shown to them and rewarded for good behaviour. Overall, 23% of all responses indicated ‘negative’ discipline methods such as ‘hit them’, ‘make them kneel down’, ‘deprive them of food’, twist their ears/pinch them’, ‘work hard’ and other forms of physical punishment, as the best methods to discipline children. Other findings included, children aged 6-14 experienced more corporal punishment than other age groups, with more boys experiencing physical punishment than girls in schools. The cane or a stick is most used to beat children. He also found that children were exploited by their school teachers for labour purposes under the guise of punishment (Kyei-Gyamfi, 2010).

The frameworks (theoretical and conceptual) outlined at the beginning of this chapter informed and guided our understanding of the key principles and processes underlying effective and constructive discipline strategies with children. These theories highlighted the importance of the complex interplay of factors that shape parental discipline responses, individual child and parent characteristics and behaviours, and more generally the contextual influences within which parenting occurs.

It is challenging to identify clear links between parenting styles, discipline strategies and child outcomes as there is lack of consensus in the definitions of strategies, methodological issues related to measuring precisely attitudes, behaviours and outcomes in question. Nevertheless, certain associations are emphasized in the literature, such as links between parental warmth, inductive discipline strategies and higher levels of moral internalization in children. Also, parental harshness and more negative behavioural outcomes for children, have been observed to play important roles in determining the pattern of effects. In general, physical discipline of children in the country is the most used form of discipline and occurs widely at the home and school.

2.14 Concluding remarks on empirical studies

Put together, most of these studies in Ghana, focused on corporal punishment with little or no mentioned of the other forms of child discipline methods. The findings of these studies were based on descriptive analysis which presented only percentages of respondent's responses without providing reasons for the observed responses and giving consideration to the clustered or hierarchical nature of the data collected. The methods used to analyze data from these surveys fall short of the best statistical approach. For example in the case of MICS 3 and MICS 4 Surveys, even though complex sample survey designs were applied in the data collection, resulting in data with hierarchical structures in the target population, only descriptive analysis of the data was provided. Consequently, simply analyzing the data that way will lead us to wrongly attribute response variation in the data and lead to misleading conclusions about the significance of the diverse sources of effect on the response. Again, it will not bring out the importance of contextual effects on responses. Such analysis do not provide in-depth information to throw more light on the issues of violence against child in the name of discipline.

The goal of the study therefore, is to re-analyze the MICS 4 survey data on child discipline using statistical models that can identify the important risk factors in child discipline practices and provide in-depth information to guide the discourse on child discipline in Ghana.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter is devoted to the methodology of the study. It also addresses the conceptual and methodological concerns associated with defining and investigating multilevel models with categorical outcomes. It describes the nature of the dataset used, operational definitions (Ghana Statistical Service, 2011) and illustrates the application of multilevel binary logit and multilevel multinomial logit models to the MICS 4 dataset in terms of how the models were derived, applied and evaluated.

3.2 The Dataset

The Dataset used in this study was derived from the multiple indicator cluster (MICS) survey 4 conducted by the Ghana Statistical Service in 2011 (Ghana Statistical Service, 2011). This study begun in 2012. The MICS dataset available then was the MICS 4.dataset. When MICS 6 was conducted in 2018, the analysis of the MICS 4 dataset was far advanced (About 70% complete). The MICS 6 dataset was not readily available. The MICS 4 dataset was therefore used for the study. The MICS 4 is one of a series of household surveys designed to provide periodic data on protective and risk factors of child development in Ghana. The survey data comprised demographic and socio-economic characteristics of households and discipline methods applied to children aged 2-14 in urban and rural households from all ten regions of Ghana.

The data were obtained through direct interviews using structured questionnaires in selected households. Parents' responses to multiple hypothetical questions involving child misbehaviour were the measure of discipline responses. Each identified household with at least a child aged 2-14 was asked 13 questions on child discipline issues (Ghana Statistical Service, 2011). One of the questions sort to find parent's support for use of physical discipline as an effective means of upbringing up children. That is, if the household supports the application of physical discipline as needed for good child upbringing. The other 12 questions were on how the household actually applied discipline to an identified child within a given reference period. Specifically, the information was obtained by reading to the head or parent or caregiver various scenarios of behaviour considered as inappropriate by the parent or caregiver in the household for the selected child and how the parent or caregiver reacted or administered punishment. That is, the behaviour or reaction of parents in response to, the perceived misbehaviour by the child (Ghana Statistical Service, 2011).

To answer the research questions, two key discrete responses in the dataset were considered; A binary response in answer to the question on whether physical discipline is needed for good child upbringing and a five level multinomial response indicating the type of method used by the household within the reference period. Nine individual household characteristics were identified. These are: sex of household head, religion of household head, educational level of household head, ethnicity of household head, and wealth index of household, number of children aged 2–14 in household, locality and region of the household.

3.3 Sampling design

The MIC 4 survey aimed to estimate indicators at the national level for urban and rural areas for all 10 regions of Ghana. A two-stage, stratified cluster sampling method was used for the selection of the survey sample. The first stage involved the selection of primary sampling units or clusters from the 2010 population and housing census sampling frame (Ghana Statistcal Service, 2010; Ghana Statistical Service, 2011) and the second stage involved the selection of sampling units or households from the selected clusters in the first stage. The data was treated as a three-level structure with households at level-1, type of area of household (rural or urban) as level-2 and region of household as level-3.

3.3.1 Sample Design and Coverage

Based on the survey design, precision and required geographical coverage, a total of 12,150 households were targeted to be interviewed in all. To achieve this, a total of 810 EAs were selected across the 10 regions of the country using the sampling frame of the 2010 population and housing census of 37,622 EAs. A two stage cluster sampling method was applied. The first stage selected 810 clusters made up of 309 urban EAs and 501 rural EAs. The second stage involved the selection of 15 households from each of the selected clusters by first sequentially numbering all listed households from 1 to n and using a systematic sampling for the selection. Systematic sampling is a probability sample selection method in which the sample is obtained by selecting every k^{th} element of the population where k is an integer greater than one. The first number of the sample is randomly selected from within the first k elements. The selection is done from an ordered list (Ghana Statistical Service, 2011).

The indicator chosen to obtain the required sample size is immunization rate consisting of full immunization, BCG, Polio 3, MMR and DPT of children age 12-23 months. To calculate the sample size several factors were considered and values for others, assumed, or taken from previous similar surveys. The factors considered in arriving at the sample size were:

r is the predicted or anticipated prevalence (coverage rate) for the indicator being estimated (expected rate for the indicator for 2011)

f is the sample deff (design effect (deff) for the indicator in 2006 MICS 3)

h is the average household size (average household size in 2006 DHS)

p is the proportion of the total population upon which the indicator, r, is based. (Proportion of children aged 12-23 months among the total population)

4 is the factor to achieve the 95% level of Confidence

$(1 + t)$ is the factor necessary to raise the sample size by 10% for non-response (t is the non response rate for households in 2006 MIC 3)

0.12r is the margin of error to be tolerated at the 95% level of confidence, defined as 12 percent of r (12% thus represents the relative error of r).

n is the required sample size, expressed as number of households, for the key indicator

The formula below was used to calculate n, the sample size

$$n = \frac{4r(1-r)f(1+t)}{(0.12r)^2hp}$$

Out of the 12,150 households initially selected, 11,925 households were successfully interviewed. The regional level response rate was more than 90 percent with the Eastern region having a response rate of 99 percent. The response rate in the rural areas was higher than that of the urban areas (Ghana Statistical Service, 2011).

After filtering out households that did not meet the criteria of at least one child in the age range of 2–14, a total of 8,156 households were left for the study (Ghana Statistical Service, 2011).

3.3.2 Background information on households

To contextualize the issues, some basic descriptive analysis were conducted to explore the distribution of household discipline practices by household characteristics. This was done to provide further insights to the study.

Table A1, in Appendix A, shows the background information of households with regards to sex of household head, region of household, residence of household (whether urban or rural), number of household members, educational level of household head, wealth index of household, religion and ethnicity of household head, age-group of household head and number of children aged 2–14 in household. The household head is generally considered as the key person that ensures the welfare of the household members.

3.4 Operational Definitions

To provide precise meaning of the concepts and variables used in the study in terms of how they were measured so that the results can properly be related to the theoretical concerns of the study, the following were defined:

Physical discipline method: A method of discipline which involves the infliction of physical pain upon a person's body as punishment (Ghana Statistical Service, 2011). In this instance, discipline is administered by an adult (as a parent or a teacher) to the body of a child ranging in severity from a slap to a spanking with the intention to cause the child to experience pain, but not injury. The purpose is for correction or control of the child behaviour (Durrant, 2005).

Non-physical discipline method: Non-physical discipline consists of both punitive and non-punitive methods, but does not include any forms of corporal punishment such as smacking or spanking Ghana Statistical Service, 2011). Non-physical discipline involves heavy parental involvement, and use of reasoning and bargaining as disciplinary methods. Non-physical discipline is also known as empathic discipline and positive discipline. It is about loving guidance, and requires parents to have a strong relationship with their child so that the child responds to gentle guidance as opposed to threats and punishment.

Psychological aggression discipline method: Psychological aggression as a discipline method refers to disciplinary practices where shouting, yelling and screaming at a child and calling a child offensive names such as 'dumb' and 'lazy are used as a way of discipline Ghana Statistical Service, 2011)..

No preference for a particular discipline method: This category consist of households whose responses indicated that they were not biased towards the usage of any particular disciplinary method. The responses of the household did not show clearly any preferred method or consistency of using a particular method of discipline. The one applying the discipline randomly applied any method on the impulse.

Table 3.1: Types of discipline methods

Type of discipline method	Description of method
Non-physical method	- Took away privileges, forbade something the child liked or did not allow child to leave house - Explain why the child's behaviour was wrong - Gave the child something else to do
Psychological Aggression method	- Shouted, yelled or screamed at child - Called the child dumb, lazy or another name - Ignored or refused to communicate
Physical method	- Shook the child - Spanked, hit or slapped the child on the bottom with bare hand - Hit the child on the bottom or elsewhere on the body with something like a belt, hairbrush, stick or other hard object - Hit or slapped the child on the hand, arm or leg - Hit or slapped the child on the face, head or ears - Beat the child up, hitting him/her all over as hard as one could
No specific method	- Household responses did not show preference to the usage of any of the above methods. Household used the above methods in equal measure or did not show bias or preference to any of the above methods.

3.5 Study Variables

3.5.1 Predictor Variables

Predictor variables are also referred to as covariate or independent variables used to explain the outcome or response variable. The predictor variables considered in this study were the socio-economic and demographic characteristics of households, where the sampled child dwells.

The variables were mostly categorical (Table D1 Appendix D). Education of household head with the levels or variable values - 'None', 'Primary', 'Middle/JSS' and 'Secondary+'. Wealth quintile of household with levels – 'Poorest', 'Second', 'Middle', 'Fourth' and 'Richest'. Age group of household head with levels – '15-21 years' , '22-40 years', '41-60 year' and '61+ years'. Number of children aged 2–14 in household with levels – '1', '2' and '3+'. Sex of household head with categories 'male' and 'female'. Location or residence of household with classification as 'rural' and 'urban'. Religion of household head with categories – 'Christian', 'Moslem', 'traditional', 'Other Religion'. Ethnic group of household head categorized as 'Akan', 'Ga/Dangme', 'Ewe', 'Guan', 'Mole-Dagbani', and 'Other'. Region of household with categories as – 'Western', 'Central', 'Greater Accra', 'Volta', 'Eastern', 'Ashanti', 'Brong Ahafo', 'Northern', 'Upper East' and 'Upper West'.

3.5.2 Response Variables

Two types of response variables or outcomes were identified. A binary response variable and a multinomial response variable. The binary response variable was obtained directly from the dataset as households responded "yes" or "no" to the question: 'Do you believe that in order to bring up, raise, or educate a child properly, the child needs to be physically punished?'

3.5.3 Identifying the multinomial response variable

Because household heads could use multiple methods of discipline, the responses were recoded to ensure that every household was identified with the *most-used* child discipline method within the survey reference period. To achieve this, a new variable was generated for each household based on the most used disciplinary method within the reference period. This was done by assessing each household's responses and based on a set condition, (the most use discipline method of household) a specific discipline method was assigned the household. The discipline methods assigned were: 'use of physical disciplinary', 'use of non-physical discipline', 'use of psychological aggression discipline', 'use of both non-physical and psychological aggression discipline' and 'use of no specific discipline method (random use of methods)'. By using this approach each household was assigned a specific disciplinary method as the most used discipline method within the reference period. The survey design did not factor-in interviewing the selected child in the household upon whom discipline was applied during the reference period.

Five categories were identified in the multinomial response outcome. Households were therefore placed in these categories consisting of 'households who mostly applied physical discipline methods', those that mostly applied 'non-physical discipline methods', those that mostly applied 'psychological aggression discipline methods', those that mostly applied 'both non-physical and psychological aggression methods' and those that 'randomly used of all methods – no specific discipline method'. The following steps were followed to construct the categories of the *most used* discipline method by the households.

1. The average 'yes' response for each category of discipline method used was obtained for each household
2. The highest average of 'yes' response for a category of discipline method used was assigned to the household as the *most used* discipline method within the reference period.
3. Households responses with no highest average of 'yes' response for any discipline method were classified in the category of random use of all methods -no specific method.

The category of physical discipline method was a combination of responses of households who indicated they used severe forms of physical discipline and other less severe forms. This combination was as a result of preliminary examination of the data set which showed there were fewer households (14%) who used this severe form of discipline method. This category was therefore collapsed and merged with the category of households that indicated they applied physical discipline. This was done so as to clarify patterns in the data and to prevent possible failure of convergence when combinations of discrete variables result in too many cells with no cases. Each household head responded to 13 questions in the child discipline module. Each household indicated the discipline methods applied and whether they believe physical discipline was needed to bring up a child properly. Table D1 in Appendix D, show all the variables used in the analysis with their variable names, type of variable and value of the variable.

3.6 Data analysis strategy

Firstly, an exploratory analysis of the data was performed to summarize and visualize the important characteristics of the dataset. These included determining relationships among the predictor variables, assessing the direction and relationships between the predictor and response variables. During the exploration, the following predictors with large categories were recoded to facilitate easy interpretation: Age of household head was regrouped from single years into 15–21, 22–40, 41–60 and 61+ years. Ethnicity of head of household was also regrouped from 11 levels into 6 major ethnic groups. Similarly, religion of head of household was regrouped from 13 levels into 4 major groups. Also, the number of children in households were regrouped into three main groups (Table D1 in Appendix D). The study design (multistage survey sampling) used to obtain the data indicates the data has multiple levels (three levels). The data structure in the population is hierarchical, and the sample data are a sample from this hierarchical population.

Secondly, a bivariate analysis was employed to identify the factors that are associated with the two outcome variables (binary and multinomial). Pearson chi-squared test and *p*-values were obtained to test the significance of each of the potential risk factors in the bivariate analysis.

Thirdly, analysis of the household responses revealed five groups of households with each using a specific discipline method or a combination of two discipline methods. There were households who applied only physical discipline methods, households that mostly applied non-physical discipline methods, households that mostly applied psychological aggression discipline methods, households that mostly applied both non-physical and psychological aggression methods and Households that randomly applied any of the methods – no specific discipline method.

Fourthly, multilevel modelling (focus of study), where the stratified nature of the dataset was taken into consideration in the study of the relationship between the potential risk factors and the outcome variables was conducted. The multilevel modelling was used to predict the outcome variable as a function of the predictors.

3.7 Developing the models

Many different approaches have been proposed to address different types of outcomes in such datasets and most importantly can answer the questions of interest. Depending on our questions of interest and structure of responses, our model was developed from the generalized linear mixed models (GLMM). Like the linear mixed models (LMMs) which allow both fixed and random effects they are particularly used when there is non-independence in the data, such as arises from a hierarchical structure. A mixed model is made up of a fixed effects part and a random effects part. For the fixed effect part, the parameters that are estimated are the coefficients of the covariates whereas for the random effect part the parameters estimated are the variances of the intercepts or slopes. The GLMM are extensions of LMMs and assume normal Gaussian random effects. The extension however, is not straight forward as it calls for the use of a link function other than the identity link function and specify an appropriate error distribution for the response at each level.

Conditional on these random effects (assumed by the Gaussian model), data can have any distribution in the exponential family. Our data has nonlinear components with hierarchical (nested) structure. With such data structures, the random effects are not necessarily normally distributed. With this in mind our model was developed considering the GLIMMIX procedure which uses two approaches. These two approaches are distinguished as "G side" random effects and "R-side" random effects. The "G side" random effects uses an approach where the random effects are included in the linear predictor. The "R-side" random effects approach is where correlation among the data is modelled directly. Again, the GLIMMIX procedure is able to fit models for non-normal data with hierarchical random effects, provided that the random effects have a normal distribution.

3.7.1 Generalized linear mix models (GLMM)

The response variable of a generalized linear model (glm) follows a distribution whose parameters depend on a linear function of covariates and is modelled by

$$\boldsymbol{Y_i} = \boldsymbol{X_i\beta} + \boldsymbol{\epsilon_i} \qquad (3.1)$$

A glm consist of three main components; a linear predictor $\eta_,$ which is a linear combination of regression coefficients $\boldsymbol{\eta_i = x_i'\beta}$, a link function $\boldsymbol{g}$ (.) that relates the mean of the response to the linear predictor, $g(\mathrm{E}([Y_i]) = \eta_i$ and a response distribution for Y_i from the exponential family of distributions eg binary, binomial, beta, gamma and inverse Gaussian distribution and normal distributions. A GLM is applied when the observations don't exhibit dependency and the data are uncorrelated. However, when there are dependencies in the observations then a GLMM is applied and this extends the class of glm by incorporating normally distributed random effects. The random effect model has the advantage of addressing both the categorical variables (dummy variables of the fixed part of model) and the random effects variables simultaneously. That is, it also has a design matrix for the random effects. Dummy variables are generally considered when dealing with single level regression whereas random effects are considered when there are multiple levels in the dataset.

A linear mixed model adds random effects to the glm model and is written as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon} \tag{3.2}$$

where

- Y is the vector of observed responses
- X is the known design matrix of fixed effects
- β is a vector of fixed (but unknown) parameters
- Z is the known design matrix for the random effects
- γ is the vector of random effects N~(0, G)
- ϵ is a vector of random errors N~(0, R)

If $Z = 0$ *and* $R = \sigma^2 I$, then the mixed model reduces to the glm

Our model is represented in the following format,

$$\text{Log (odds)} = \beta_0 + \beta_{1j}X_{1j} + \beta_{2j}X_{2j} \dots \dots, +\beta_{nj}X_{nj} \tag{3.3}$$

The model under consideration is presented in the format, where β refers to the parameters and X represents the predictor variables.

- β_0 is the intercept
- β_{ij} are the slopes that indicate the direction and strength of association between X_s and the response variable
- X_{ij} are the characteristics of the predictors

The log (odds) or log-odds ratio is defined as: $\ln\left(\frac{p}{1-p}\right)$ and is expressed as the natural logarithm of the ratio between the probabilities that an event will occur $P(y = 1)$ to the probability that it will not occur. The predicted probability that an event will occur is defined as;

$$P = \frac{1}{1 + exp^{-z}} \tag{3.4}$$

$where \quad z = \beta_0 + \beta_{1j}X_{1j} + \beta_{2j}X_{2j} + \cdots \dots \dots, +\beta_n X_{nj}$

3.7.2 Multilevel models

Multilevel models depicts the hierarchical structure of a dataset and thus makes it easier to perceive the model in stages. Hierarchical structured dataset, lends itself naturally to the application of multilevel modelling. In such data structures the observed characteristics are handled as the explanatory or predictor variables, and the unobserved characteristics are addressed by the random effects. The random components in multilevel models indicate the presence of between-unit variation and within-unit correlation. The multilevel modelling approach therefore facilitates identification of the effects of covariates and quantify variation in the data by using both fixed and random effects.

A hierarchical structured dataset with three-levels can be modelled with a three-level model that has fixed effects at the first and second levels and random intercepts and slopes at the second and third levels with $i, j \ and \ k$ denoting the indices of level-1, level-2 and level-3 units respectively. Such a model can be applied to the MICS 4 dataset which can similarly be specified in three stages or levels. At each stage or level, covariates and random effects are incorporated to explain the level-specific variation around the mean intercept and mean slope. The level-1 model suggests a linear relationship between the fundamental observed response Y_{ijk} and the level-1 covariate x_{ijk} and can be expressed as;

$$y_{ijk} = \beta_{0jk} + \beta_{1jk}x_{ijk} + e_{ijk} \tag{3.5}$$

At level-2, the intercept and slope from the level-1 model vary among level-2 units according to the following relationships with the level-2 covariate x_{jk} ;

$$\beta_{0jk} = \pi_{00k} + \pi_{01k}x_{jk} + \gamma_{0jk} \tag{3.6}$$

$$\beta_{1jk} = \pi_{10k} + \pi_{11k}x_{jk} + \gamma_{1jk} \tag{3.7}$$

Finally, the level-2 intercepts vary among level-3 units according to the level-3 models;

$$\pi_{00k} = \tau_{00} + \gamma_{0k} \tag{3.8}$$

$$\pi_{10k} = \tau_{10} + \gamma_{1k} \tag{3.9}$$

In addition to the responses, covariates, and the parameters that relate them, this three-level model incorporates random terms at each of the levels with level-1 residual e_{ijk} and random-effects vectors at level 2 and level 3, and $\gamma_{jk} = (\gamma_{0jk}, \gamma_{1jk})$ and $\gamma_k = (\gamma_{0k}, \gamma_{1k})$, respectively. The distribution of the random effects is assumed to be a normal distribution;

$$\gamma_{jk} \sim N(0, G^2) \quad and \qquad \gamma_k \sim N(0, G^3)$$

The covariance matrices G^2 and G^3 specify how the random intercept and slope vary across the level-2 and level-3 units respectively, with the residual vector of the level-3 units being $e_k \sim N(0, R^3)$.

In the context of our study the dataset has individual households (level -1) nested in location or type of area (level-2) and further nested in region (level-3) of household.

To address the research questions with the dataset, two key models were considered (the binary logit model and multinomial logit model) based on the type of responses from the households.

The logit and the probit models

Two popular regression models for dichotomous data are the logit model (logistic regression) and the probit model (probit regression). The logit and the probit models are types of glms and have the form$f(u_x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$. Both models can be used to model the relationship between one or more numerical or categorical predictor variables and a categorical outcome. They both solve the problems of the regression line predicting outside the range of 0 and 1, by taking any number and rescaling it to fall between 0 and 1. They take the linear model and feed it through a function to yield a nonlinear relationship.

For large datasets, the probit model fits marginally better but it makes no difference in conclusions. The only difference is in the link functions they use. Whereas the logit model uses the cumulative distribution function of the logistic distribution ($Pr\ (Y = 1|X) = (1 + e^{-X'\beta})^{-1}$), the probit model uses the cumulative function of the standard normal distribution ($Pr\ (Y = 1|X) = \Phi(X'\beta)$). In this study, we considered the covariates as directly connected to the probability of success, so we choose the logistic regression because it uses a canonical link and appears to have an easier interpretation compared to the probit model.

3.7.3 The Binary logit model

The binary logit model has dependent variables that are restricted to two categories (dichotomous). The binary logit model is appropriate for modelling dichotomous dependent variables compared to using the OLS approach where the regression line may lead to prediction outside the range of zero and one and residuals plot would reveal heteroscedasticity. The binary logit model addresses these issues by fitting a nonlinear function to the data. That is, it replaces the linear straight line of OLS with a sigmoidal (S shape relationship curve) that respects the boundaries of the dependent variable and assumes proper specification of independent variables that eliminates heteroscedasticity.

The model allows for clustering (non-independence) when fitting a multilevel model with group-level random effects. The study focused on showing how multilevel models can be applied to three-level binary response data that allowed for correlation between responses of households in the same type of area or region, and to explore the extent of between-locations variation in responses to use of physical discipline as necessary.

Considering a three-level data structure where a total of n households (at level 1) are nested within J groups (at level 2), n_{jk} households in group j and by n_k in group k. We used 'group' as a general term for any level unit, e.g. type of area (location) or region.

By denoting y_{ijk} as the response for household *i* in group *j*, and group k, and x_{ijk} at the household level covariate; we obtain:

$$y_{ijk} = \beta_0 + v_k + \mu_{jk} + e_{ijk} \tag{3.10}$$

where the levels 2 and 3 residuals v_k, u_{jk} and the level 1 residuals e_{ijk} are assumed to be independent and follow the normal distributions with zero means:

$$v_k \sim N(0, \sigma_v^2) \quad , \quad \mu_{jk} \sim N(0, \sigma_v^2), \qquad e_{ijk} \sim N(0, \sigma_v^2) \tag{3.11}$$

This model can also be expressed in terms of the mean or expected value of y_{ijk} for a household *i* in group *j and group k* and with value x_{ijk} on x.

For a binary response y_{ijk}, the $\mathrm{E}(y_{ijk} | x_{ijk}, u_{jk}) = \pi_{ijk} = \Pr(y_{ijk} = 1)$ and a generalized linear random intercept model for the dependency of the response probability π_{ijk} on x_{ijk} is written:

$$F^{-1}(\pi_{ijk}) = \beta_0 + \beta_1 x_{ijk} + \cdots \ldots.. \tag{3.12}$$

where F^{-1} ("*F* inverse") is the link function, taken to be the inverse cumulative distribution function of a known distribution.

F^{-1} is the link function because it links π_{ijk} to the covariates. The link function is simply a function of the mean of the response variable *Y* that we use as the response instead of *Y* itself. The selection of the link function is based on the type of the outcome variable and its distribution in the data sample, the characteristics of the model being developed, and the metric to use to report the results. In our instance the scale of measurement for the first outcome variable is dichotomous (binary) variable. Consequently, a binomial sampling distribution and the logit link function will be applied. The logit link function is the natural log of the odds that Y equals one of the categories.

Random intercept logit model

In a logit model, $F^{-1}(\pi_{ijk})$ is the log-odds that $y = 1$ and equation (3.12) becomes

$$\log\left(\frac{\pi_{ijk}}{1-\pi_{ijk}}\right) = \beta_0 + \beta_1 x_{ijk} + u_{jk} \tag{3.13}$$

where β_0 is interpreted as the log-odds that $y = 1 \; when \; x = 0 \; and \; u = 0$ and is referred to as the overall intercept in the linear relationship between the log-odds and x. The exponential of β_0, exp (β_0), gives the *odds* that y = 1 for *x* = 0 and *u* = 0.

In analyzing the dataset, the interest was in the amount of variation that could be attributed to the different levels in the data structure and the extent to which variation at a given level can be explained by the covariates. The variance partition coefficient measures the proportion of the total variance that is due to differences between groups.

Predicted response probabilities

For more than one level, the expression for the response probability is obtained as

$$\hat{\pi}_{ijk} = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_{ijk} + \hat{u}_{jk})}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_{ijk} + \hat{u}_{jk})} \tag{3.14}$$

For an individual *i* in group *j nested in group k by* substituting the estimates of $\beta_0, \; \beta_1,$ $and \; u_{jk}$ obtained from the fitted model.

3.7.4 The Multinomial logit model

The second key response variable in the dataset has a multinomial outcome where the outcome has more than two possible events that are not ordered. In such a situation, multinomial logit model is an appropriate analytic procedure. This model is similar to the Binary model except that the dependent variable is not restricted to two categories. Because the outcome represents a probability between 0 and 1, a linear regression would not be appropriate as it would result in predictions that fall outside the allowable range of the dependent variable. The probability distribution of the multinomial logit model therefore represents an extension of the Bernoulli model for dichotomous outcomes. The difference is that the outcome consists of more than two unordered categories, which are separately compared to a selected reference category. For the multinomial model outcome, we assume that the probability distribution is multinomial and we use a logit link function.

Denoting the total number of outcome categorization to C with each individual category indexed by c such that the probability of being in the *cth* outcome category $P\ (Y = c)\ is\ \pi_c$, where $c = 1, 2, \ldots, C.$. The cumulative probability of each possible outcome $\pi_1, \pi_2, \ldots, \pi_c$ can be expressed such that their sum is one.

One of the categories is use as a reference category so there will actually be $C - 1$ equations to be estimated. The probability of membership in one of the other categories is compared against the probability of being in the reference category.

For individual i, then, the probability of being in category c $(c = 1, 2, \ldots, C - 1)$ versus the reference group (C) can be defined as follows:

$$\eta_c = \log\left(\frac{\pi_1}{\pi_c}\right) \tag{3.15}$$

In general, $c = 1, 2\ldots, C - 1$ logits for the categories can be defined:

$$\log\left(\frac{\pi_1}{\pi_C}\right), \log\left(\frac{\pi_2}{\pi_C}\right), \ldots \ldots, \log\left(\frac{\pi_{C-1}}{\pi_C}\right), \quad (3.16)$$

If there are only two categories, this will be the same as the dichotomous case considered in section 3.7.3. The outcome scores are transformed using the cumulative logit link function into an unobserved (or latent) continuous variable η_c that describes the log odds of being in a particular category c in contrast to a reference category with the variance;

$$Var\ (Y_c|\pi_c = \pi_c(1 - \pi_c) \quad (3.17)$$

In the dataset the outcome of the multinomial comprised five unordered response categories as follows;

1. Households that mostly applied physical discipline methods,
2. Households that mostly applied non-physical discipline methods,
3. Households that mostly applied psychological aggression discipline methods,
4. Households that mostly applied both non-physical and psychological aggression methods
5. Households that randomly applied any of methods – no specific discipline method.

Table B1 in Appendix B, shows the percentage distribution of the categories as follows: Non-physical discipline methods category - 23.0%, Psychological aggression methods category - 26.4%, Physical discipline methods category - 7.7%, Non-physical and psychological aggression methods category - 31.9% and use of random discipline methods category - 11.0%.

Developing the multilevel multinomial model

The multinomial logistic regression model was used to predict the odds of individual i in group j being in outcome category c relative to outcome C (reference category) using the set of q predictors. We defined

$$\eta_{cij} = \log\left(\frac{\pi_{cij}}{\pi_{Cij}}\right) = \beta_{0jc} + \beta_{1jc}X_{1ij} + \beta_{2jc}X_{2ij} + \cdots + \beta_{qjc}X_{qij} \quad (3.18)$$

where there is no separate residual variance term at Level 1 because the variance is dependent upon the mean. The Level 1 variance was again set to a scale factor of 1. Equation (3.18) was summarized to link the expected values of the outcome to the predicted values of η_{cij} as follows:

$$\eta_{cij} = \log\left(\frac{\pi_{cij}}{\pi_{Cij}}\right) = \beta_{0j(c)} + \sum_{q=1}^{Q} \beta_{qj(c)}X_{qij} \quad (3.19)$$

At level 1, continuous variate η_{cij}, is a ratio of two odds (i.e., the probability of each category c versus the selected reference category C that is explained by a set of linear combination of X predictors ($q = 1 \ldots. Q$)

At level 2, the model is expressed as

$$\beta_{qj(c)} = \gamma_{q0(c)} + \sum_{s=1}^{Sq} \gamma_{qs(c)}W_{sj} + u_{qj(c)} \quad (3.20)$$

At this level, one or more level 1 intercepts or slopes are modelled as a function of a set of level 2 predictors (*W*) and corresponding variance terms (u_{qj}), keeping in mind that for each intercept or slope modeled, there will be *C* – 1 equations.

Applying the model to the dataset

An unconditional (no predictors) model was first estimated to examine the extent of variability of the nominal outcome across the region of household. This provided at level 1, *C* – 1 estimates for individual households *i* in region *j* as follows:

$$\begin{aligned}
\eta_{1ij} &= \log\left(\frac{\pi_{1ij}}{\pi_{Cij}}\right) = \beta_{0j(1)} \\
\eta_{2ij} &= \log\left(\frac{\pi_{2ij}}{\pi_{Cij}}\right) = \beta_{0j(2)} \\
\eta_{3ij} &= \log\left(\frac{\pi_{3ij}}{\pi_{Cij}}\right) = \beta_{0j(3)} \\
\eta_{4ij} &= \log\left(\frac{\pi_{4ij}}{\pi_{Cij}}\right) = \beta_{0j(4)}
\end{aligned} \tag{3.21}$$

At level 2, the combined set of the models suggest that the intercepts vary between the regions as follows:

$$\begin{aligned}
\beta_{0j(1)} &= \gamma_{00(1)} + u_{0j(1)} \\
\beta_{0j(2)} &= \gamma_{00(2)} + u_{0j(2)} \\
\beta_{0j(3)} &= \gamma_{00(1)} + u_{0j(3)} \\
\beta_{0j(4)} &= \gamma_{00(4)} + u_{0j(4)}
\end{aligned} \tag{3.22}$$

In all, four parameters were estimated across the two models

Computing predicted probabilities

When there are more than two categories for the dependent variable, the probabilities for the outcomes are computed differently. For a dependent variable with C categories, $C - 1$ log odds equations are calculated. To do this, the estimated intercepts for each category was calculated using equation (3.23) to obtain the likelihood or the predicted probability of belonging to each of the discipline method categories.

$$\pi_{cij} = \frac{\exp(\eta_{cij})}{1 + \sum_{c=1}^{C-1} \exp(\eta_{cij})} \tag{3.23}$$

where η_{cij} is the value of the linear component for specific values of the predictors. For the reference category the value of η_{cij} is 0 and $exp(0) = 1$, the log odds are therefore 0 [log(1) = 0], and the odds ratio is 1 ($e^0 = 1$). The probability of being in the reference category is then

$$\pi_{cij} = \frac{1}{1 + \sum_{c=1}^{C-1} \exp(\eta_{cij})} \tag{3.24}$$

3.8 Testing for whether a multilevel model is needed

This test was done to establish if a multilevel model fits the data significantly better than a single-level model with no higher cluster or cluster effects. To conduct this test, firstly, a single level model was fitted to the data using maximum likelihood to estimate the parameters that most likely characterize the data. Secondly, a three level model was also fitted to the data using the maximum likelihood. Thirdly, a likelihood ratio test was then used to compare the fits of the single level model and the three level model. The null hypothesis that there are no higher level effects was tested by comparing the three-level model to a single-level model.

For the model with equation;

$$y_{ijk} = \beta_{0jk} + e_{ijk} \qquad \text{and} \qquad e_{ijk} \sim N(0, \sigma_e^2) \tag{3.25}$$

The null and alternative joint hypotheses;

$$H_0: \sigma_v^2 = 0, \quad \sigma_u^2 = 0 \ \ (no\ location\ and\ no\ regional\ variation)$$

$$H_0: \sigma_v^2 > 0,\ \sigma_u^2 > 0 \ \ (\ significant\ location\ and/or\ regional\ variation)$$

The joint hypothesis related to the variances of the random effects was tested using the likelihood ratio (LR = (ΔG^2) test. The LR test statistic (ΔG^2) for testing the null joint hypothesis was calculated as;

$$\Delta G^2 = (-2logL_0) - (-2logL_1) \tag{3.26}$$

where $L_0\ and\ L_1$ are the likelihood values for the single-level model and the three-level model respectively and a p-value of $P(x_k^2 >= \Delta G^2)$. Large values of ΔG^2 lead to small p-values, which provide evidence against the single level model and in favour of the multilevel model. The LR test value was then compared to a chi-squared distribution with 9 degrees of freedom equal to the number of extra parameters in the model. The resulting p-value was used to assess whether the three-level model fits the data significantly better than the single-level model.

3.9 Interpretation of parameters

3.9.1 Variance components

The variance components of a multilevel model can be interpreted in three ways using the coverage intervals (CI), variance partition coefficients (VPC) and intraclass correlation coefficients (ICCs). The CI gives the absolute size of variance components in the metric of the response variable and their relative magnitudes. The VPC statistics indicates the proportion of the response variance at each of the model's hierarchy (i.e. measure of the proportion of the total variance that is due to differences between groups), and the ICC statistics measure the expected degree of similarity (or homogeneity) between responses within a given cluster.

3.9.2 Coefficients and predicted probabilities

Suppose there are *n* individuals in group *j*, denoting y_{ij} the nominal response for individual *i* in group *j* and x_{ij}as an individual-level explanatory variable, the multilevel random intercept can be expressed as

$$log\left(\frac{\pi_{ci}}{\pi_{1i}}\right) = \beta_{0c} + \beta_{1c}x_i. \qquad c = 2, \ldots.. M \tag{3.27}$$

In the random intercept model equation (3.15), β_{0c} is interpreted as the log-odds that an individual with $x_{ij} = 0$ and $u_{cj} = 0$ will be in category c rather than category 1 of the response y. The log-odds of being in response category c rather than category 1 will vary across groups, according to the variance of u_{cj}, and group with $u_{cj} = 0$ and will be located at the mean of this distribution. The parameter β_{1i} is the effect of a 1-unit change in x on the log-odds of response category c versus response category 1 after adjusting for group effects.

The random effects u_{cj} allows the response probabilities to vary across groups. Specifically, if $u_{cj} > 0$ it indicates that a household in group *j* has above-average chance of being in response category c rather than category 1. On the other hand if $u_{cj} < 0$ than we would expect the ratio of π_{cij} to π_{1ij} to be lower than average for households in groups i and j.

For a dependent variable with *C* categories, we calculate *C* – 1 log odds equations taking one of the categories as the reference group. To calculate the probability for each category, we take each one of the *C* – 1 log odds computed and then exponentiate it and divide each by the sum of the odds to obtain the probability for each category. The predicted probabilities of the response outcome c for individual households i in group j was therefore calculated from the fitted model as:

$$\hat{\pi}_{cij} = \frac{\exp(\hat{\beta}_{0c}+\hat{\beta}_{1c}x_{ij}+\hat{u}_{cj})}{1+\exp(\hat{\beta}_{0c}+\hat{\beta}_{1c}x_{ij}+\hat{u}_{cj})} \qquad c = 2, \ldots., M \qquad (3.28)$$

The intercept β_{0c} for the contrast c is the log of the probability of being in category c relative to the probability of being in category 1 when $x = 0$, and $\exp(\beta_{0c})$ is the ratio of the probability of being in category c to the probability of being in category 1. The coefficient of x for the contrast c, (β_{1c},) is the effect of a 1-unit increase in x on the log-odds of being in category c rather than category 1. Predicted response probabilities were calculated from the estimated coefficients for the values of x. These were substituted in the estimates for $\beta_{0c}, \beta_{1c}, \beta_{2c,}\beta_{3c}$ from the fitted model.

3.10. Significance testing

To test for a relationship between x and the log-odds of being in category c rather than category 1, a Z-test of the null hypotheses that the coefficient of x, (β_{1c},

$\beta_{2c,}\ \beta_{3c}$) is zero was conducted. The test statistic (Z-ratio) was calculated as the estimate of β_{1c} divided by its standard error and compared with a standard normal distribution.

3.11 Contextual effects

These are covariates defined at levels other than level 1. The coefficient of a level 2 covariate is called a contextual effect. In multilevel models they include cross-level interactions between a level 1 and a higher level covariate which allow the effect of the level 1 to depend on the value of the higher level variable.

3.12 Correspondence between observed and predicted probabilities

To measure the performance of probability predictions for the binary responses (as a result of differences between the observed and predicted probabilities), the Receiver Operating Characteristics (ROC) curve was used to measure the classification performance of the binary model. For the nominal responses (multinomial model) a confusion matrix was created to assess the performance of the model in terms of correct classification. That is, the model's ability to successfully assign a specific outcome to each case.

3.13 Covariate effect on the response variable (fitting the model)

Estimates of parameters from the multinomial model (choice of discipline method) and the binary model (support or approval for use of physical discipline) were obtained to assess their effects on the response variables of the models.

3.14 Principles of multilevel modelling

When the outcome of interest of a study and its observed and unobserved determinants have a hierarchical structure, where factors that influence the outcome arise from a variety of levels of aggregation or observation, than one can safely apply multilevel modelling. For example, where units (level 1) are nested within groups (or clusters – level 2) and the groups are themselves nested within super groups (level 3), the data form a three-level hierarchy and three-level models can be fitted to account for the additional level.

When various levels of structures arise in a dataset leading to higher levels of clusters that significantly differ from one another, on the outcome then it becomes important to incorporate them in the study. Simply fitting a two-level model to a three-level data will lead to attributing wrong response variation to the two included levels (Moerbeek, 2004; Tranmer & Steel, 2001; Van et al., 2009) which may in turn lead to drawing misleading conclusions about the relative importance of the different sources of influence on the outcome.

The outcome of interest with multilevel modelling takes place at the individual level referred to as the lower-level outcome. In analyses with more than two levels, this is called the level-1 outcome. At the lower level, individual outcomes are usually influenced in part by individual and micro-level characteristics. Again, the characteristics of a location have similar effects on all individuals residing within that location. The varying levels of outcomes and determinants, i.e., at the individual level, location level or regional level, gives rise to the hierarchies analyzed in multilevel models. In such situations, the resulting data are clustered with hierarchical structure and expressed in terms of correlation among the measurements on units within the same cluster. Statistical models for clustered data must account for the intra-cluster correlation (at each level) failure of which can result in misleading inferences. In applying a multilevel model it is important to decide which parameter values will be fixed or random and which estimation type (ie maximum likelihood estimation or restricted maximum likelihood estimation) would be used (Fidell Barbara & Tabachnick Linda, 2007).

3.14.1 Non-normal outcomes and multilevel modelling

Regression models for the analysis of categorical dependent variables are generalizations of logistic regression analysis to polytomous data. Important questions to address in such analysis include; Is the outcome - discrete? continuous? binary?, or multinomial?. The analysis in this study used a three level multilevel model with non-normal outcomes (hierarchical generalized linear models - dichotomous and polytomous) to answer the research questions. The hierarchical generalized linear models are suitable for data which are categorical, has non-normally distributed response variables including binary, proportional, counts or ordinal data. For this type of model, the assumptions of normal distribution and homoscedastic errors are violated (Hox, 2010; O'Connell, et al., 2008).

To address these problems, the outcome variable is transformed using a nonlinear link function. Also, an appropriate non-normal error distribution is needed to be incorporated in the model so that the model building strategies and interpretations will be applicable (Luke, 2004). For multilevel models with dichotomous outcomes, the binomial distribution (i.e., Bernoulli) and the logit link are used for estimation of the parameters. Similarly, the models with polytomous, ordinal-type outcomes use a multinomial distribution and a cumulative logit link for the computations. In formulating and testing these models consideration was first given to the fully unconditional model where no predictors are included in the model and secondly the conditional model where predictors are introduced into the model.

3.15 Building the models to answer the research questions

3.15.1 Building the multilevel binary response model

To obtain the best fitting model for the dataset and use it to answer the research questions, the following model building strategy (Raudenbush & Bryk, 2002) was adopted:

1. An unconditional model containing no covariates but random effects for the intercepts was first estimated
2. Results of this first model was used to compute the intraclass correlation coefficient (ICC) which estimates how much variation in the response variable exists between location (urban/rural) and region units.
3. Additional models were estimated by introducing additional parameters while checking the fit statistics for improvement in the model fit after each model is estimated.

Specifically, the following steps were followed in building and identifying the best model to address the research questions for the binary response model.

Table 3.2: Model building steps for the Binary logit model

Model	Effects added
1	Model with no predictors. Just random effect intercepts for levels 2 and 3
2	Model with predictors of level 1 and random effect intercepts of levels 2 and 3
3	Model with predictors of levels 1 and 2 and random effect intercepts of levels 2 and 3
4	Model with levels 1 and 2 predictors and random effects intercept for level 3

3.15.2 Building the multilevel multinomial response model

The multilevel multinomial logistic regression model was considered to predict the probability of being in one of the discipline categories and the effects of the predictors. Since the outcome is nominal, the logit link function was considered. The following steps were used to build the multinomial model;

The unconditional model M1 with no predictors was first determined. This was followed by the conditional model M2 (M1 plus three predictors). Conditional model M3 (M2 plus two additional predictors) and finally a conditional model M4 made up of model M3 plus two more predictors.

3.16 Model evaluation and diagnostics

When models are developed it is important to evaluate and measure how accurately they describe the outcome variable in the sample and in the population. A goodness of fit or calibration of a model, tests whether the model adequately describes the variations in the response variable by investigating how close values predicted by the model are to the actual observed values.

Model evaluation and diagnostics involve performing, goodness of fit test (Likelihood ratio test to provide a better fit to the data), statistics tests for individual predictors (Wald test- to evaluate the statistical significance of each coefficient in the model) and validation of predicted values (*k*-fold cross validation) using simulations.

Primarily, there are three methods of validation; split sample validation, cross validation and bootstrapping validation. These methods are considered as internal validation methods and they involve data-splitting, repeated data-splitting, jack-knife technique and bootstrapping.

The split sample validation involves randomly splitting the data into two samples (training sample and validation sample). There are also model validation metrics like the Kolmogorov-Smirnov (KS) test measures which check whether the model is able to separate events and non-events. Other validation metrics include rank ordering where the percentage of events in each decile group is calculated and checked to determine whether the event rate is monotonically decreasing. Another metric is the Area Under Curve (AUC) which explains the trade-off between the true positive rate (sensitivity) and true negative rate (specificity). The Hosmer Lemeshow test is another model validation metric which measures calibration and shows how close the predicted probabilities are to the actual rate of events, i.e the degree of agreement between the observed probability and the predicted probability.

As a general rule, “when too small a portion of the available data is reserved for model fitting, the ability of the model to predict future observations suffers” (Picard & Berk, 1990). It is therefore very important to appropriately split the data for validation, taking into consideration the available data.

Another approach to the assessment of predictive accuracy of models, is the use of external validation methods where the model is applied to a totally new data that was not part of the model development. This approach is considered the most rigorous and unbiased test for the model and the entire data collection process. The idea here involves first excluding a sub sample of observations, secondly developing a model based on the remaining sample, and thirdly testing the model on the originally excluded sample.

To evaluate the performance of the developed models, the following methods were applied: The *k*-fold cross validation, the bootstrap method.and model classification.

3.16.1 The *k*-fold cross validation

The *k*-fold cross validation has to do with finding how well a model performs in predicting the target variable on different subsets of the data. This approach involves randomly dividing the set of observation into k groups, or folds, of approximately equal sizes. The first fold is treated as a validation set, and the method is fit on the remaining *k-1* folds to predict the target variable in the testing data. The process is repeated *k* times, with the performance of each model in predicting the hold-out set being tracked using a performance metric such as accuracy. That is, to use a limited sample in order to estimate how the model is expected to perform in general when used to make predictions on data not used during the training of the model.

Importantly, each observation in the data sample is assigned to an individual group and stays in that group for the duration of the procedure. This means that each sample is given the opportunity to be used in the hold out set one time and used to train the model *k-1* times. The *k* value must be chosen carefully to be representative of the data sample. It must be chosen such that each train or test group of the data samples is large enough to be statistically representative of the broader dataset.

The steps followed to perform a 10-fold cross validation on the study dataset are as follows:

1. The binary model first fitted with the study dataset (using the significant predictors) to obtain the; maximum likelihood estimates of the models' coefficients, the overall significance of the models and the partial significance of each variable included in the model. Also obtained was the association table of predicted probabilities and observed responses.
2. The "c" statistics (a measure of discrimination of the binary model) in the generated association table gives the area under curve (AUC) of the (ROC) curve. This indicates the percentage of how well the model performed.
3. A number of ROC curves were then generated to calculate the area under the curves for further assessment.
4. Also calculated were the sensitivity and specificity of the binary model.

3.16.2 The Bootstrap method (Simulating to evaluate a model's performance)

The Bootstrap method is a nonparametric technique that simulate data directly from a sample. In such simulations the concern is with the variability in the estimated values of the model's parameters over *n* repetitions. If the variability is large, then the model's coefficients are highly dependent on the specific portion of the original data used to fit the model. It also means that the model is not stable (an indication of over fitting). In other words, if the distribution of the estimates of the model's parameters obtained through simulation on a sample of the data, do not average around the same values of the estimates computed on the whole original sample, then there is internal instability of the developed model.

The Bootstrap technique is often used when no statistical model of the data is evident. The method use the fact that the empirical cumulative distribution function (CDF) of the data is an approximation to the underlying distribution. Generally, to implement the method, you first fit a parametric distribution to the data using the maximum likelihood estimator (MLE) and then you sample from this parametric distribution. The central assumption for bootstrapping is that the original sample accurately represents the actual population. Specifically, three main steps are required: Compute the statistic of interest on the original data, then resample n number of times from the data to form n bootstrap samples and finally compute the statistics on each resample.

In applying the bootstrap method to assess the developed model for the study, the following questions were answered: (1) How close are the estimated parameters to the values of the true parameter? (2) Does the 95% Confidence Interval for the estimated parameters include the true parameter values? Also, the standard errors of estimated parameters were compared as the number of observations increased. Finally, the Wald intervals with the empirical 95% quintiles of the sampling distribution were compared with increasing sample size.

3.16.3 Model Classification (Confusion Matrix)

Model classification (Confusion matrix) is a matrix used for evaluation of statistical models. It describes the model's ability to identify the existing classes in a given dataset by comparing the actual to predicted values for each predicted state.

To use this method, a cross classification Table (Confusion matrix) of the actual and predicted response variables for the dataset used to fit the model was created. Table 4.15 shows the confusion matrix by cross classifying the actual and predicted response variables. Table 4.15 also shows when the actual and predicted response levels agree. The 'mean' match, which is the proportion of observations correctly classified by the model is also indicated. We compare the results of the original dataset (null model) and the results when the developed model is applied for the assessment.

3.16.4 Categorical predictor variables in regression analysis

Categorical variables have a measurement scale consisting of a set of categories. They have two main types of measurement scales; natural ordering where the categorical variables have ordered scales and are called ordinal variables. When they have unordered scales, they are called nominal variables. For nominal variables the order of listing of the categories is irrelevant. Statistical methods designed for nominal variables give the same results no matter how the categories are listed. This cannot be said of ordinal variables where reordering the categories any how could affect the results.

Categorical variables used as predictors in regression analysis requires special attention. This is because regression analysis requires numerical variables. Unlike dichotomous or continuous variables, categorical variables cannot be entered into the regression equation just as they are. To ensure that the results are interpretable, the categorical variables to be included should first be recoded into a set of separate binary variables called dummy coding and subsequent creation of a table called contrast matrix. This is done automatically by SAS.

In general, a categorical variable with k levels will be transformed into k-1 variables each with two levels. For instance if a categorical variable has six levels (k=6), five dichotomous variables (design variables) would be constructed that would contain the same information as the single categorical variable. This way, the dichotomous variables, could be entered directly into the regression model.

For instance, the study data has educational level of household head as a categorical variable with four levels (1 = None, 2 = Primary, 3 = Middle/JSS, 4 = Secondary +). In this instance there will be three new variables. For all, but one of the levels of the categorical variable, a new variable will be created that has a value of 1 for each observation at that level and 0 for all others. In the study data X_1 (None) will have a value of 1 for each observation in the dependent variable and 0 for all other observations. Likewise X_2 would be 1 when a household head has primary level education and 0 otherwise, and X_3 would be 1 when the household head has middle/JSS and 0 otherwise. The level of the categorical variable that is coded as 0 in all the new variables is the reference level or the level to which all of the other levels are compared to. The educational level categorical variable, Secondary+, is the reference level in the study data.

3.17 Summary on methodology

This chapter provided an overview of the methodology used in the development of the study models using multilevel modelling. It introduced the concepts of the study with their precise meaning and relationship to the theoretical concerns that motivated the study. The chapter also covered the assumptions of multilevel models, when to use them, and how to deal with hierarchical models with discrete outcomes. The mathematical equations behind multilevel models, interpretation of parameters, significance testing and contextual effects were also discussed. The chapter ended with discussions on the processes used in developing the models to answer the research questions and model evaluation and diagnostics.

CHAPTER 4

RESULTS AND DISCUSSION

4. 1 Introduction

This chapter presents the detailed results, interpretation of estimated parameters and discussions in terms of the study's objectives and key responses. It starts with analysis of the background information and bivariate analysis of the independent and response variables. The chapter also provides information on the fitted models and interpretation of the content of the tables generated by the model results and answers to the research questions. Also presented in this chapter, are the evaluation results (classification results of the models) and discussions on the robustness of the developed models. The statistical packages SPSS and SAS 9.4 were used in the analysis.

4.2 Analysis of background information

Table A1, in Appendix A, shows the descriptive analysis of the background information. Table A1, shows that, a little over half of household heads (50.3%) approve the use of physical discipline for proper child upbringing. The analysis indicates that, the Eastern (68.1%) region, had the highest proportion of household heads who approved the use of physical discipline. Household heads with no education (51.5 %) tended to accept the use of physical discipline compared to household heads with higher education (42.2%). The Ga/Dangme ethnic group had the highest proportion of household heads (57.7%) who approved applying physical discipline. There is no significant difference between the proportion of male and female household heads (50.4 % and 50.0% respectively) who approved the application of physical discipline.

Household heads in the age category 61 and beyond, have the highest proportion (51.2%) who approved the application of physical discipline.

4.2.1 Bivariate analysis of Objective 1 (Binary response variable)

Table A2 in Appendix A, presents the bivariate analysis of the predictor variables by the level of the response variable. The bivariate analysis showed that for all households who approved, the application of physical discipline, the Mole/Dagbani ethnic group had the highest proportion (18.5%). The analysis further showed there is an association between the ethnicity of household head and support for use of physical discipline is significant. ($\chi^2 = 60.32$, $p < 0.0001$). Again, for households at the lowest quintile (poorest households) there was not much difference in their approval or disapproval of the use of physical discipline. At the highest quintile level (richest households), however, we have a higher proportion who disapproved (6.3%) compared to those who approved (4.1%). The Pearson chi-square test indicates the association between the wealth index and the response variable is significant (.$\chi^2 = 52.17$, $p < 0.0001$).

For households with one child, the proportion who indicated their disapproval for use of physical discipline (15.7%) was higher compared to those who approved (13.2%). A higher proportion of households with three or more children (23.4%) approved the use of physical discipline compared to households who disapproved (21.2%). Pearson chi-square test indicates a significant association between the two variables (number of children and approval for use of physical discipline). Also, religion of household head, has a significant association ($p<0.0001$) with approval for use of physical discipline whilst the age group of household head indicates a marginal significance (p<0.0243).

4.2.2 Bivariate analysis of Objective 2 (multinomial response variable)

This section presents the bivariate results of the second key response variable (method of discipline used by household). Table B2 in Appendix B presents a cross-tabulation of region of household by type of discipline method applied by household. The last column shows the total households in the study sample from each region along with the percentages of households who applied a particular discipline method.

Overall, 23.0% of households applied non-physical methods, 26.4% applied psychological methods, 7.7% applied physical methods, 31.9% applied a combination of non-physical and psychological aggression methods, 11.0% applied a combination of all methods (random use of all methods). We observed that Northern region had the highest respondents (19.8%) in the sample. Of the 1,615 respondents in this region, 4.3% applied non-physical methods, 6.2% applied psychological aggression methods, 2.1% applied physical methods, 6.3% applied non-physical and psychological aggression methods and 0.9% randomly applied all methods.

The region with the lowest proportion of respondents is the Eastern (5.6%) region where the most applied discipline method by households was a combination of non-physical and psychological aggression methods. Pearson's chi-square test for association showed that the region of household was associated with the type of discipline method applied ($\chi^2 = 540.30$, $p < 0.0001$). The null hypothesis that the two variables were independent was rejected and the conclusion made that there is association between them. In other words, the type of disciplinary method used by households varies significantly across the regions.

4.2.3 Predictors by type of discipline method

In terms of level of education of household head and the type of discipline method applied, the results show that (Table B3) household heads with no education were in the majority in the application of all the different types of discipline methods (non-physical methods - 12.2%, psychological aggression methods - 16.9%, physical methods - 5.4%, non-physical and psychological aggression methods - 17.3%) with the exception of usage of all methods (random use of discipline methods) where household heads with Middle/JSS were highest (6.2%). Pearson chi-square test for association indicates that, the education level of household head is associated with the type of discipline method applied ($\chi^2 = 1164$, $p < 0.0001$). This means the use of the different discipline methods varies significantly at the different levels of education of the household head.

Similarly, in terms of ethnic groupings, Mole Dagbani household heads are in the majority (3.2%) of those who applied physical discipline methods followed by Akan household heads (1.4%). Pearson chi-square test for ethnicity ($\chi^2 = 358.85$, $p < 0.0001$) indicates a significant association between ethnicity and type of discipline method applied.

Households in the lowest wealth index quintile have the highest proportion of heads who applied physical discipline (3.7%) with households in the highest wealth index quintile having the lowest proportion of heads who applied physical discipline. A

Pearson chi-square test ($\chi^2 = 230.64$, $p < 0.0001$) indicated the association between the two variables was significant.

Table B3 in Appendix B, again shows the proportion of male headed households (5.9%) who applied physical discipline methods was higher than the proportion of female headed households who applied physical discipline (1.8%). The Pearson test of association between gender and type of discipline however, was not significant ($\chi^2 = 4.64$, $p < 0.3258$).

Table B3 in Appendix B further shows that the association between religion of household head and type of discipline method is significant ($\chi^2 = 212.07$, $p < 0.0001$). Also, the age group of household head has a significant association with the type of discipline method applied in the household. Again, the number of children aged 2–14 in a household has a significant association with the type of discipline method applied in the household.

The results of the bivariate descriptive analysis show that nearly all the predictors indicated significant association with the response variables. Five predictor variables were significant under objective 1 and 6 predictor variables significant under objective 2.

4.2.4 Exploratory Data analysis

The goal here was to understand the essential features of the distribution of the values of the variables which comprised the dataset and to identify the important ways in which the variables appear to be related. The exploration focused on statistical measures which pertained to the central tendency, dispersion, and shape of the data distribution.

The exploration analysis confirmed the distribution pattern of both the binary responses/outcome and multinomial outcome that they were not normally distributed. This was expected because the outcome/response variables were both categorical in nature. It further addressed the following concerns before the actual modelling: sample structure identification, solving the problem of missing data, identification of relevant predictor variables (see Appendix C for the exploratory data analysis, tables and charts).

The charts in Appendix C show the histograms, probability plots and boxplots for the variables indicating how their values in the data spread out and their distributions. For example, the variable *physical punishment needed,* the values for the data appear uniformly distributed on the interval (1, 2). The mean and the median appear identical (2.8 and 3.0), and the distribution has a skewness of positive 0.06. The negative kurtosis measure (-1.42) indicates that the tails of this distribution are lighter than for a normal distribution. The exploratory analysis further showed the distribution of a number of the study variables are non-symmetric, and their kurtosis measures are negative indicating more data points clustered at their upper ends than at the lower ends of their distribution. With the exception of the variable *number of children in household age 2-14*, all other variables are positively skewed indicating that their data values located to the right of the means are more spread out than data values to the left of their mean. The positive kurtosis measure of the variable *ethnicity*, denotes heavy tails of its distribution. Relationships between the variables were also examined to measure the strength of relationships (correlation coefficient) among the variables.

4.3 Testing for whether multilevel model is needed

To establish whether or not a multilevel model fits the data significantly better than a single-level model which includes no higher cluster or cluster effects at all, a maximum likelihood, and likelihood ratio tests were conducted to compare their fit statistics. The null hypothesis that there are no higher level effects at all was tested by comparing the three-level model to a single-level model. The test uses the ratio of the maximized value of the likelihood function for the full model (L_1) over the maximized value of the likelihood function for the simpler model (L_0). The likelihood ratio test statistic is defined as:

$$-2\log\left(\frac{L_0}{L_1}\right) = -2[\log(L_0) - \log(L_1)] = -2(L_0 - L_1)$$

Table 4.1 shows the likelihood ratio statistic results for the single level (1 level) model and the multilevel (3 levels) model.

Table 4.1: Fit Statistics of the single model and multilevel model

Models Fit Statistics	Single level model	Multilevel model
-2 Log Likelihood	10,945.88	10,904.43
AIC	10,965.88	10,926.43
AICC	10,965.91	10,926.47
BIC	10,983.02	10,929.76
CAIC	10,993.02	10,940.76
HQIC	10,972.12	10,922.78

Table 4.1, shows the -2 log-likelihood value for the multilevel model is 10,904.43, compared to 10,945.88 for the single-level model. This gives the difference between the two likelihood ratio test statistics as (LR):$\Delta G^2 = 10945.88 - 10904.43 = 41.45$. This difference (41.45), compared with a chi-squared critical value for the test at the 5% level with 9 degrees of freedom is 16.92, indicating the difference is significant.

The LR test therefore shows that a multilevel approach to analyze the data is clearly favoured over a single-level approach. The null hypothesis of no contextual differences was therefore rejected and conclusion that there is evidence of between-level variation in the responses made. Rejection of the null hypothesis implies that there are group differences, in which case the multilevel model is preferred over the single-level model.

The dataset for the study have two key set of responses on child discipline practices. The first set of responses is a binary response where respondents had to response ‘yes’ or ‘no’ to a question. The second set of response had to do with household use or application of a type of discipline method. The multilevel binary logit model was used to address the first part of the responses because the dataset was hierarchical with two categorical responses. Section 4.4 presents’ results and discussion of the multilevel binary logit model and section 4.5 presents’ results and discussion of the multilevel multinomial responses.

4.4 Results and Discussion of Objective 1

This section discusses the results of Objective 1 of the research by answering research questions 1 to 3 using the fitted binary logit model described in Section 3.7.3

Table 4. 2: Model information of the Binary logit outcome

Response Distribution	-	Binary
Response Variable	-	Physical discipline needed
Link Function	-	Logit
Variance Function	-	Default
Variance Matrix Blocked By	-	Region of household
Estimation Technique	-	Maximum Likelihood
Likelihood Approximation	-	Laplace
Degrees of Freedom Method	-	Containment
Number of Observations Read	-	8156
Number of Observations Used	-	8156
Response Profile		
‘no’ physical discipline not good	-	No
‘yes’ physical discipline is good	-	Yes
Total Frequency		
No	-	4053
Yes	-	4103

Table 4.2 shows information about the fitted binary logit model. The information include the link and variance functions, the distribution of the response and information on the estimator and computational algorithm used for the estimation theory.

Furthermore, it shows the number of observations read from the input data and sum of frequencies read and used. The response profile indicates the ordered value for the binary data, the probability being modelled and the reference category for the model (SAS Reference Materials, SAS 9.4). The multilevel binary logit model described in Section 3.7.3 above was constructed for this response variable using all the selected covariates.

4.4.1 Level effects on the response variable

Table 4.3: Parameter estimates of the multilevel Binary logit model

Effect	Estimate (s.e)	DF	t-value	Pr > |t|	Alpha	Lower	Upper
Intercept β_{000}	-0.0850 (0.08)	40	-1.03	0.31	0.05	-0.25	0.08

Cov Parm	Subject	Estimate	Standard Error	Z -Value	Pr > Z
Intercept μ_{jk}	Region	0.2308	0.06	3.75	<0.0001
Intercept v_k	Location (Region)	0.0061	0.02	0.91	0.25
e_{ijk}	level 1 error variance	3.29			

e_{ijk} = level 1 error variance (Snijders & Bosker, 1999, O'Connel et al., 2008)

Table 4.3 shows that the overall mean response for the binary logit model (β_{000}) is estimated to be -0.0850 (Unconditional model). The total variance calculated as the sum of the three variance components $v_k + \mu_{jk} + e_{ijk}$ is estimated to be $0.0061 + 0.2308 + 3.29 = 3.5269$ with the level 1 error variance (e_{ijk}= 3.29). The variance of a logistic distribution with scale factor 1.0 is $\pi^2/3$, or approximately 3.29 (Hedeker & Mermelstein, 2007). The total variance is made up of the variance within the type of area location v_k (estimated to be 0.0061) and the variance between the different regions (estimated to be 0.2308). To interpret these three variance components, the correlation coefficients, variance partition coefficients and the intraclass correlation coefficients were considered.

Table 4.4: Estimates of Variance Components of fitted Binary logit model

Level	95% Coverage Interval	Variance Partition Component (VPC)	Intraclass correlation coefficient (ICC)
	(-1.96σ, +1.96σ)	$\text{Var}(y_{ijk}) = \text{Var}(\sigma_e^2) + \text{Var}(\sigma_l^2) + \text{Var}(\sigma_r^2)$ $\sigma_e^2 + \sigma_l^2 + \sigma_r^2$	
Location	$(-1.96\sigma_l, +1.96\sigma_l)$ $(-0.1530,\ 0.1530)$	$\frac{\sigma^2_{\text{location}}}{\sigma_e^2 + \sigma^2_{\text{loation}} + \sigma^2_{\text{region}}}$ $= \frac{0.0061}{3.29+0.0061+0.2308}$ $= 0.0017$	$\text{ICC}_{\text{location}} = \frac{\sigma^2_{\text{location}}}{\sigma_e^2 + \sigma^2_{\text{loation}} + \sigma^2_{\text{region}}}$ $= \frac{0.0061}{3.29 + 0.0061 + 0.23082232}$ $= 0.0017$
Region	$(-1.96\sigma_r, +1.96\sigma_r)$ $(-0.9416,\ 0.9416)$	$\frac{\sigma^2_{\text{region}}}{\sigma_e^2 + \sigma^2_{\text{loation}} + \sigma^2_{\text{region}}}$ $= \frac{0.2308}{3.29+0.0061+0.2308}$ $= 0.0654$	$\text{ICC}_{\text{region}} = \frac{\sigma^2_{\text{region}}}{\sigma_e^2 + \sigma^2_{\text{loation}} + \sigma^2_{\text{region}}}$ $= \frac{0.2308}{3.29 + 0.0061 + 0.2308}$ =0.0654
Household	$(-1.96\sigma_e, +1.96\sigma_e)$ $(-3.5551,\ 3.5551)$	3.29	3.29

Table 4.4 shows the estimates of the three variance components (coverage interval, variance partition and intraclass correlation coefficient) of the fitted model. The 95% coverage interval for the variance component is the range of response values within which we expect 95% of the associated random effects to lie. That is, given the overall mean response (β_{000} = -0.0850), and variance between the different regions (σ^2_{region} =

0.2308), we say that 95% of the regional effects of 'yes' responses or approval of physical discipline are expected to lie in the range of - 1.0267 to 0.8566 which is calculated as $[(-0.0850 - 1.96\sqrt{0.2308})$ and $-0.0850 + 1.96\sqrt{0.2308}).]$ respectively. Using this calculation our results show that there is substantially meaningful variation in 'yes' responses at the household and regional levels of the model. At the location level (type of area), however, the results show no meaningful variation (0.0017 ~ 00%).

The Variance Partition Component Statistics (VPC) shows the relative magnitude of the variance components. The results indicate that at the location level (type of area), there is very little variation (almost zero percent variation) in the 'yes' responses. At the regional level, however, 6.54% variation lies between the regions. The remaining variation of 93.5 % $(1 - 0.0017 - 0.0654 = 0.9329)$ is attributed to the household level.

The ICC measures the expected degree of similarity or homogeneity between the responses within a given cluster or level. The strength of the correlation depends on the extent to which two households belong to the same higher level units. Households who share the same level, will be more alike and therefore more correlated than households from different levels. The greater the differences between levels, the more similar households will appear within their levels.

In summary, the 95% coverage intervals, the VPC statistics and the ICC statistics show that there is some degree of clustering and significant variability in the intercepts across the regions in the data. This situation justifies developing a multilevel model to answer the research questions.

Table 4.5: Summary Table - Parameter estimates for fitted Binary logit models

Effects		**Model 1** No predictors specified. Only levels 2 and 3 random effect intercepts	**Model 2** Four level 1 predictors specified, plus level 3 random effect intercept	**Model 3** Model 2 plus all 8 level 1 predictors specified and level 3 random effects
Intercept		-0.0850(0.08)	0.2799(0.36)	-0.4121(0.34)
Fixed Effects				
Education of Head	Middle/JSS		0.1870 (0 .25)	-0.0072(0.26)
	None		0.3334 (0.25)	0.08078(0.26)
	Primary		0.2009 (0.25)	-0.0866(0.26)
	Secondary +		0	0
Ethnicity of Head	Akan		0.5798 (0.26)	0.6635*(0.26)
	Ewe		0.4445 (0.27)	0.492(0.27)
	Ga/Dangme		0.8015 (0.28)	0.8419*(0.29)
	Gruma		0.5331(0.27)	0.5271*(0.27)
	Grusi		0.4736 (0.26)	0.4887(0.26)
	Guan		-0.2592 (0.27)	-0.2377(0.28)
	Mande		0.4894 (0.30)	0.488(0.30)
	Mole Dagbani		0.6719 (0.25)	0.685*(0.25)
	Non-Ghanaian		0.2054(0.31)	0.2545(0.31)
	Other		0	0
Sex of Head	Female		0.0053(0.06)	0.0521(0.06)
	Male		0	0
Religion of Head	Christian		-0.1972(0.06)	-0.1597(0.10)
	Moslem		0.0143(0.11)	0.07253(0.11)
	Traditional		-0.0887(0.11)	-0.0841(0.11)
	other		0	0
Age Group of Head	15-21 years			-1.0438*(0.37)
	22-40 years			0.0593(0.07)
	41-60 years			-0.0084(0.06)
	61+years			0
Children aged 2-14 in household	1 child			-0.2305*(0.6)
	2 children			0.0311(0.06)
	3+			0
Wealth index of household	Fourth			-0.2124*(0.09)
	Middle			-0.0466(0.08)
	Poorest			0.0557(0.07)
	Richest			-0.5165*(0.10)
	Second			0

CONTINUATION OF TABLE 4.6

Random Effects				
Covariance Parameter				
Level-3 (Region)	Intercept	0.2308**(0.06)	0.2677**(0.07)	0.2667**(0.07)
Level-2 location (Region)	intercept	0.0061(0.02)		
Model fit				
-2Log (CD|r.effects)		10,873.87	10,788.12	10,718.35
-2Log likelihood		10,993.28	10,910.53	10,842.47
AIC		10,997.28	10,946.53	10,896.47

Note: *p < 0.05; **=likelihood ratio test significant; Values based on SAS PROC GLIMMIX. Entries show parameter estimates with standard errors in parentheses; Estimation method = Laplace; Best fitting model; CD=Physical Discipline is appropriate for good child upbringing

Table 4.5 presents the summary of all the parameter estimates and standard errors in brackets for the fitted binary models. Model M1 is the unconditional model with no covariates and models M2 and model M3, the conditional models with the introduction of the covariates. The best model is the model with the minimum objective function (-2LL) value. The changes in these values (-2LL) across models M1 to M3 indicate that the inclusion of the covariates improved the fit considerably. Model M3 with the least objective function (-2Log Likelihood (CD|r.effects)) value of 10,718.35 was considered the best fit for the model and therefore used for the analysis. Model M3 showed the estimate of the variance component of region (level-3) intercept as 0.2667 which is more than twice its standard error of 0.07. This again, confirms the evidence of significant variation in the 'yes' responses across the regions. The results again showed that there is virtually no variation (0.00% ICC) in the 'yes' responses across locations (urban, rural divide).

Table 4.5 answers research question 2 (under Objective 1) on whether there is variability in households' responses in the approval of physical discipline for child upbringing. Table 4.5 shows that the estimate of the variance component at the regional level (level-3) intercept is 0.2308 (M1, the unconditional model) with a standard error of 0.06. This estimate (0.2308) is more than twice its standard error (0.06), and suggests there is evidence of significant variation in the responses across the regions. The results further show that there is no variation in the responses across locations (urban, rural divide). This result also confirms the ICC value for location level (type of area).

This finding of variability of intercepts across regions, indicates that support for use of physical discipline varied significantly across regional levels. This means, the regional groupings appear to influence a household's perspective and attitude to approve or support the use of physical discipline in child upbringing. There could be some specific characteristics of the regions i.e., ethnic diversity, socio-cultural practices, etc., that interact to influence household heads to support or otherwise the application of physical discipline. It would be interesting to identify what specific characteristics of these regions that influence households to support or approve the use of physical discipline as a better way to bringing up children.

Specifically, our findings indicate that households in regions like Brong Ahafo, Eastern, Northern, Upper East, Volta, and Western have a higher likelihood of favouring the use of physical discipline compared to Ashanti, Central, Greater Accra, and Upper West regions.

Size of level effects

Using the predicted probabilities of responding 'yes', under the assumption that the group effects (levels 2 and 3 effects) follow a normal distribution, we would expect approximately 95% of group level units to have values within two standard deviations of the mean of zero. These will give approximate coverage intervals of:

Group level 2 (location/type of area): $\pm 2\sqrt{\sigma^2{}_{location}} = \pm 2\sqrt{0.0061} = \pm 2(0.0781) = \pm 0.1562 \ (-0.1562 \ to + 0.1562\ )$.

Group level 3 (Region): $\pm 2\sqrt{\sigma^2{}_{region}} = \pm 2\sqrt{0.2308} = \pm 2(0.4804) = -0.9608 \ to + 0.9608$.

Substituting the values of the estimated of intercept (-0.0850) from Table 4.5, we obtain the following predictions for group level 2 (location).

For 2 standard deviations below the mean:

$$P_{"yes"} = \emptyset_{ijk} = \frac{e^{y_{ijk}}}{1+e^{y_{ijk}}} = \frac{e^{(-0.0850-0.1562)}}{1+e^{(-0.0850-0.1562)}} = \frac{0.7856}{1.7856} = 0.4399$$

and for 2 standard deviations above the mean:

$$P_{"yes"} = \emptyset_{ijk} = \frac{e^{y_{ijk}}}{1+e^{y_{ijk}}} = \frac{e^{(-0.0850+0.1562)}}{1+e^{(-0.0850+0.1562)}} = \frac{1.0737}{2.0737} = 0.5178$$

This means that the proportion responding 'yes' at the group level 2 (location-type of area) will lie between 0.4399 and 0.5178 in the middle of 95% of group level 2 households. Similarly the proportion responding 'yes' at the group level 3 (regional) will lie between 0.2600 and 0.7059 in the middle of 95% of group level 3 households as follows;

$$P_{"yes"} = \emptyset_{ijk} = \frac{e^{y_{ijk}}}{1+e^{y_{ijk}}} = \frac{e^{(-0.0850-0.9608)}}{1+e^{(-0.0850-0.9608)}} = \frac{0.3514}{1.3514} = 0.2600$$

$$P_{"yes"} = \emptyset_{ijk} = \frac{e^{y_{ijk}}}{1+e^{y_{ijk}}} = \frac{e^{(-0.0850+0.9608)}}{1+e^{(-0.0850+0.9608)}} = \frac{2.4007}{3.4007} = 0.7059$$

The narrower confidence interval (0.4399 to 0.5178) for the location level, means that there is very little or no variability in the responses. This is due to the positive correlation in the responses at this level and therefore resulting in no level effect. At the regional level, however, because of variability in the responses the confidence interval is wider (0.2600 to 0.7059).

These figures mean that when we ignore clustering we underestimate the variability in the data which results from the positive correlation in the data indicated by the narrower confidence interval. These findings demonstrate the dangers of ignoring data clustering and highlights the point that when clustering is ignored there is a high chance of reporting significant differences between groups when no such differences actually exists.

4.4.2 Predictor effects on the response variable and discussion of objective 1 results

Table 4.5 (Table A3 in Appendix A) shows estimated parameters for the fixed effects of the final model. The categorical variables in the model were first recoded into a set of separate binary variables called dummy variables and a contrast matrix table before obtaining the estimated coefficients for each level of the categorical variable. Table 4.5, for example shows the p-values (Pr >|t|) for hypothesis testing for all the educational levels of head of household which indicates that the estimated coefficients of the different levels of the educational predictor variable are not statistically significant or not different from 0 at the 95% confidence level. This means, the education level of head of household has no effect on household's approval or disapproval to use of physical discipline in child upbringing.

For the categorical predictor variable, age-group of household head, the coefficient of level 15 – 21 years was negative (-1.0438) and statistically significant (p <0.0048). This means that within this age group, a parent is less likely to approve use of physical discipline than when a parent is in the 61+ category (reference category). In other words, being in this age group is associated with a lower predicted log odds of approving the use of physical discipline relative to being in the reference age group (61+). The coefficients of the categorical predictor variable, ethnicity of household head, were positive and significant for Akan households heads (0.6635, p = 0.0118), Ga/Dangme household heads (0.8419, p = 0.0031), Gruma household heads (0.8419, p = 0.0482) and mole Dagbani (0.6850, p = 0.0060). These also mean that household heads from these ethnic groups are more likely to approve the use of physical discipline than their counterparts in the reference category (Other). That is, these household heads are associated with having a higher predicted log odds of approving the use of physical discipline compared to being in the reference category (Others). The gender of the household head is represented by the categorical variable Sex. The categorical variable sex of household head, was statistically not significant. The coefficient of the first level (1 child) of the categorical predictor variable, number of children aged 2-14, was negative and significant (-0.2124, p <0.0001) meaning households with one child are more likely not to use physical discipline than when there are three or more children aged 2-14 in the household. In other words, if a household has only one child aged 2-14, that household will be associated with a lower predicted log odds of approving use of physical discipline relative to the reference group (i.e.3 or more children aged 2-14 in household). For the categorical predictor variable, wealth index of household, the coefficients of the 'richest' level (-0.5165, $p < 0.0001$) and 'fourth' level (-0.2124, $p = 0.0138$), were negative and significant. These results also indicate that, households in these levels of wealth index are more likely not to approve the use of physical discipline compared to households in the reference category ('second'). Such households would be associated with lower predicted log odds of approving the use of physical discipline relative to the reference category level 'second'.

Generally, it is perceived that less-educated parents are those who tend to use corporal punishment (Durrant, 1999). This observation, however, is not conclusive, as other studies also link violent disciplinary tactics to higher levels of education (Ghana Statistical Service, 2011). The finding of this study collaborates with findings which suggest that one's educational level does not have effect on one's approval or disapproval to use physical discipline on children (Wolfner & Gelles, 1993). Paquette et al. (2000), observed that the characteristics of the larger family in which the parent and child belong to, do have influence on the likelihood that the parent may or may not support the application of physical discipline. As family size increased, support for the use of physical disciplinary methods by parents increased. The finding from this study, however, suggests that, the household size per se is not the issue, but rather the number of children aged 2–14 in a household. In other words, when the number of children aged 2-14 is large in a household, the chances are that the household heads would approve the use of physical discipline methods. This observation could also be related to the challenges and demands associated with being a parent in such a household, and the responsibility of taking care of a large number of children at home, especially if one also has to contend with the stresses of a low socio-economic status.

Research findings have shown that the age group of the household head has some relationship with the decision to approve or not to approve the application of physical discipline in bringing up children. Dietz (2000) and Giles-Sims et al. (1995), found that younger parents aged 18–29 tend to hit their children more than older parents, suggesting an association between youthfulness and violence. This study, however, found that younger parents' age 15-21 years relative to the reference group (61+) tended not to approve the use of physical discipline in child upbringing. A possible reason could be their thinking that physical discipline could ruin the relationship between them and their only child and make their child feel rejected by them and also lead to the child fearing and avoiding them.

The acquisition and possession of certain assets by households can be used to gauge the wealth index of a household and linked to the behaviour and altitude of parents, in terms of discipline, towards their children. Wealthy households may be more knowledgeable about parenting methods because of the high probability of access to resource materials on discipline from books, the media etc. Poor households, on the other hand, can be confronted with a myriad of challenges both at home and in the environment that could stress them and lead to the approval of use of force to vent out their frustration. "Wealthier households therefore may not resort to violent disciplinary practices as often" (McLoyd & Jayaratne, 1994). From the analysis of this study, a significant negative relationship between the wealth index (richest and fourth levels compared to the reference level of households and their approval for the use of physical discipline was observed. The observation is in agreement with the conclusion of Youssef et al. (1998), who also found that as the socio-economic status of families declined, rates of parental use of corporal punishment rose.

"Parenting values and beliefs have influence on the discipline responses of parents to their children" (Pinderhughes et al., 2000). From our findings, it appears there is something in the traditions of these ethnic groups (Akan, Ga/Dangme, Gruma and Mole Dagbani) that impact on household heads from these ethnic groups to approve the use physical discipline methods on children. Finding out what exactly is causing this would be an interesting further research.

Straus and Mathur (1996), found that mothers favour physical discipline less compared to fathers. They observed that most male household heads in the Ghanaian society are expected to administer discipline on serious matters in the home involving child misbehaviour. "Such discipline is mostly left for the male parent to handle when he arrives at home and a report is made to him by the wife. The male parent, therefore, is seen as the disciplinarian, in spite of the fact that mothers spend more time with the children than the father". This observation, however, was not supported by this study, as the finding indicate no inclination of the gender of the head of household towards the use of physical discipline.

Studies show that, the extent to which families perceived and apply discipline (normativeness), varies across countries. In other words, when parents regard a discipline method as normative, they feel more confident and justified in using it, and thus may be more likely to apply it carefully rather than in an impulsive and unregulated manner. In some cultures, there are forms of hostile behaviour towards children, such as spanking and shouting, which are both legally and culturally normative. However, with the increasing recognition of the problem of child maltreatment, more severe forms of aggression, such as hitting the child with a belt, etc., although still legal, have become less acceptable. Lansford et al. (2005), also found that "cultural normativeness is lowest in Thailand, China, the Philippines, and Italy and highest in Kenya, with varying collectivist and religious affiliations among these countries".

From our binary outcome model, the evidence uncovered suggest that, nearly half of Ghanaian household heads approve the use of physical discipline as a proper way to bring up children. This is a worrying observation that calls for public education to dispel this belief. Parental child discipline practices, varies across the regions, but not between the urban and rural areas. The fitted model indicates that as a household's circumstance improve in terms of its wealth index, the likelihood to apply physical discipline decreases. Households with younger parents are less likely to approve use of physical discipline whilst an increasing number of children aged 2–14 in the household leads to higher chances of the household approving the use of physical discipline. Household characteristics such as gender, educational level of head and religion, have no statistical significant influence on parent's attitudes and practices in terms of supporting the use of physical discipline.

From the dataset, the frequency of 'yes' response for use of physical discipline is 4,103 and that for 'no' response is 4,053. This gives an observed sample probabilities of 4,103 / 8,156 = 0.5030 and 4,053/8,156 = 0.4969 for 'yes' and 'no' responses respectively.

From the fitted model, the predicted response probability for a 'yes' response was obtained as;

$$P_{"yes"} = \emptyset_{ijk} = \frac{e^{y_{ijk}}}{1+e^{y_{ijk}}} = \frac{e^{-0.0850}}{1+e^{(-0.0850)}} = \frac{0.9185}{1.9185} = 0.4787$$

$$P_{"no"} = 1 - \emptyset_{ijk} = 1 - 0.4787 = 0.5213,$$

where - 0.0850 is the estimated intercept.

This answers the first research question under Objective 1 on the probability of a household approving physical discipline for child upbringing. From the calculations, the predicted probability that a household would approve the use of physical discipline is 0.48. This rate is considered high when considering the potential for corporal punishment to escalate into abuse (Straus et al., 2014).

The use of physical discipline has been found to enhance aggression, delinquency, and antisocial behaviour in children, as well as being associated with reducing the quality of relationship between parent and child, and a higher risk of being a victim of physical abuse. It is also known to increase aggression, criminal and antisocial behaviour in adults, and make it more likely that victims abuse their own child or spouse (Gershoff, 2002; Grogan-Kaylor & Otis, 2007; Holden, 2015a; Smith et al., 2005). Physical punishments like spanking teaches children that it is okay to hurt people when angry and this can lead the receiving child to believe you solve problems by hitting. Such children may continue this way of thinking into adulthood causing them to hit their spouse or children and the cycle continues.

The observed high support for the application of physical discipline in Ghana places children in the country at a high risk of experiencing the serious consequences thereof. This result, also observed by other researchers (Deater-Deckard et al., 2003; Twum-Danso, 2010), indicates the larger cultural context of the acceptability of physical discipline and the lack of debate and awareness of its potential negative outcomes and resulting legal issues. A lot more effort is therefore needed to address the notion that use of physical discipline is acceptable for proper child upbringing. As it currently stands, if nothing is done, the situation can also lead to intergenerational transmission of physical discipline methods as the norm in Ghana.

The predicted probability of a 'yes' response (0.4787), however, differs from the observed sample probability of a 'yes' response (0.5030). This difference could be that an important covariate or predictor has been left out or the mathematical form of the model could be a poor approximation of reality. It could also be that a covariate that has gone in the model as a linear effect, in fact acts non-linearly. To assess whether this difference could affect the robustness of the developed model, a model evaluation and diagnostics was performed using ROC analysis to ascertain the models performance (section 4.6)

Table A4 in Appendix A shows the odd ratio (OR) estimates of the predictor variables. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure. In other words how to quantify the effects on the response outcome as a result of changes in the exploratory variable. For categorical exploratory variables, the odds ratio compares the odds for the outcome between one level of the explanatory variable and the reference level. Considering our significant categorical predictor variables: ethnicity of head, age group of head, and number of children aged 2-14 in household and wealth index of household, the table shows that Ewe household heads (OR = 0.842, 95% CI: 0.683, 1.039) are about 16% (1-0.842 = 0.158 ~ 0.16) less likely to approve of physical discipline compared to Akan household heads (reference group). Similarly, Gruma (OR=0.872, 95% CI: 0.678, 1.122) and Grusi (OR=0.840, 95% CI: 0.643, 1.097) household heads are about 13% and 16% respectively less likely to approve of physical discipline compared to Akan household heads. Again, Guan (OR=0.406, 95% CI: 0.309, 0.534) and Non-Ghanaian (OR=0.664, 95% CI: 0.448, 0.985) household heads are 59% and 34% respectively less likely to approve of physical discipline compared to Akan household heads.

Table A4 further shows that household heads in the age group 22-40 years (OR=3.014, 95% CI: 1.464, 6.201) are about 3 times more likely to approve of physical discipline compared to household heads in age group 15-21 years (reference group). Similarly, household heads in age group 41-60 years (OR=2.816, 95% CI: 1.369, 5.793) and 61+ years (OR=2.840, 95% CI: 1.374, 5.868) are both about 2.8 times more likely to approve of physical discipline compared to their counterpart in age group 15-21 years.

Table A4 again shows that households with two (OR=1.299) or more children (OR = 1.259) aged 2-14 years are 1.3 times more likely to approve of physical discipline methods. Compared to households in the fourth quintile level (reference group) of wealth index, households in the middle quintile level (OR = 1.180) are 1.2 times more likely to approve of physical discipline methods. Households in the poorest quintile level (OR = 1.307) and second quintile level (OR = 1.237) are 1.3 and 1.2 times (respectively) more likely to approve use of physical discipline methods. On the other hand, households in the richest quintile (OR = 0.738) of wealth index are 0.7 times less likely to approve us of physical discipline methods.

This analyses answers research question 3 under Objective 1 on whether household characteristics do influence households approval or otherwise of the use of physical discipline.

4.4.3 Summary results and discussion of Objective 1

The results show that the predicted probability that a household head would approve of physical discipline as appropriate for child upbringing in the overall sample **is 0.4787.** At the rural/urban level, an ICC of 0.0017, indicates no significant influence of the level on household responses. The regional hierarchy level of the data indicates approximately 6.54% (ICC of 0.0654) of the variability in the data is accounted for by the region level. Household heads in the age group 15-21 were more likely not to approve use of physical discipline compared to age group 61+ (reference group). In terms of approving use of physical discipline, Akan, Ga/Dangme, Gruma and Mole Dagbani household heads were more likely to approve use of physical discipline methods.

Households with only one child aged 2-14, are less likely to approve use of physical discipline compared to households with two or more children in that age group. Households in the richest category of wealth index are less likely to approve use of physical discipline methods compared to households in the second category of wealth index.

4.5 Results and Discussions of Objective 2

This section of the results discusses Objective 2 of the research by answering research questions 4 to 8 using the multilevel multinomial logit regression model described in Section 3.7.2. Table B1 in Appendix B presents information about the fitted multilevel multinomial logit model for the dataset: these include, the link function, the distribution of the responses, the estimation methodology and the likelihood approximation. Also, displayed are the number of observations read from the input data and the sum of frequencies read.

4.5.1 The unconditional multinomial logit model results

Four models were estimated in this analysis. The first was fitting an unconditional model to determine the probability of households applying a specific discipline method (Research question 4 under Objective 2). The second model examined the effects of the covariates (Sex of household head, age-group of household head and number of children aged 2-14 in household) on the response variable (types of discipline method used by household), the third model examined the effects of religion, ethnicity and the educational level of household heads, and the fourth model examined the effects of the wealth index of the household including the fixed effects of two levels of random effect intercept on the multinomial outcome. In the multinomial analysis the location level (urban/rural) was dropped because in the binary analysis it indicated no significant influence on the response variable or produced any variation in the responses. The multinomial outcome analysis, therefore, considered only two levels i.e., the household and regional levels.

4.5.2 Predicted probability of choice of discipline methods

This section addresses the fourth research question on the probability of a household choosing a particulate discipline method or belonging to each of the response categories. Table 4.6 shows the number of households' *n* that applied the different types of discipline methods and the computed observed response probability for application of each method.

Table 4.7: Computed observed response probabilities for the multinomial logit model

k	Type of discipline method	n	Observed sample response probability $\pi_k = \frac{n}{8156}$
1	All methods (random use of methods)	899	0.1102
2	Non-physical and psychological discipline methods	2,602	0.3190
3	Non-physical methods	1,877	0.2301
4	Physical discipline methods	625	**0.0766**
5	Psychological methods	2,153	0.2639
	Total	8,156	1.000

Table 4.8: Fit Statistics for multinomial unconditional distribution model at α=0.05

Effect	Discipline types	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper
Intercept (k=2)	Non-physical and Psychological methods	-0.0592	0.5395	8052	-0.94	0.3453	-1.567	0.5483
Intercept (k=3)	Non-physical methods	-0.8921	0.5598	8052	-1.59	0.1111	-1.990	0.2054
Intercept (k=4)	Physical methods	-2.4622	0.7162	8052	-3.43	0.0006	-3.866	-1.057
Intercept (k=5)	Psychological methods	-0.5588	0.5732	8052	-0.942	0.3297	-1.683	0.5649

Using the parameter estimates (intercepts) from the null (unconditional) model (Table 4.7, page 128), and substituting in equation (3.23) the predicted probability (PP) for each response category of the four estimated models with just the intercept were computed as follows:

PP to apply non-physical and psychological aggression discipline methods; Category 2,

$$P(Y=2) \quad = \frac{\exp(\eta_{cij})}{1+\sum_{c=1}^{C-1}\exp(\eta_{cij})}$$

$$= \frac{\exp(-0.5092)}{1+\exp(-0.5092)+\exp(-0.8921)+\exp(-2.4622)+\exp(-0.5588)}$$

$$= 0.2257$$

PP to apply non-physical discipline methods; Category 3,

$$P(Y=3) \quad = \frac{\exp(\eta_{cij})}{1+\sum_{c=1}^{C-1}\exp(\eta_{cij})}$$

$$= \frac{\exp(-0.8921)}{1+\exp(-0.5092)+\exp(-0.8921)+\exp(-2.4622)+\exp(-0.5588)}$$

$$= 0.1512$$

PP to apply physical discipline methods; Category 4

$$P(Y=4) \quad = \frac{\exp(\eta_{cij})}{1+\sum_{c=1}^{C-1}\exp(\eta_{cij})}$$

$$= \frac{\exp(-2.4622)}{1+\exp(-0.5092)+\exp(-0.8921)+\exp(-2.4622)+\exp(-0.5588)}$$

$$= 0.0320$$

PP to apply psychological aggression discipline methods; Category 5

$$P(Y=5) \quad = \frac{\exp(\eta_{cij})}{1+\sum_{c=1}^{C-1}\exp(\eta_{cij})}$$

$$= \frac{\exp(-0.5588)}{1+\exp(-0.5092)+\exp(-0.8921)+\exp(-2.4622)+\exp(-0.5588)}$$

$$= 0.2151$$

PP for random use of all methods (no specific discipline method) reference category 1

$$P(Y=1) \quad = \frac{1}{1+\sum_{c-1}^{C-1}\exp\left(\eta_{cij}\right)}$$

$$= 0.3761$$

This can also be obtained as: $1-0.2257-0.1512-0.0320-0.2151=0.3760$

With the exception of the calculated predicted probability value for *random use of all methods* which was higher (0.3761), than the observed probability (0.1102), all the predicted probability for the other discipline methods were lower than their calculated observed response probabilities. This indicates the average response rate in the data is different from the average predicted probability. To assess whether these differences could affect the robustness of the model a classification table was constructed (section 4.7.1) to verify this observation.

Choice of discipline methods

The results show that, the most likely discipline method a household would use for correction of child misbehaviour is a combination of non-physical and psychological aggression discipline methods (Probability = 0.23) in contrast to the reference category (random methods). It is interesting to note that despite the high proportion of household heads who indicated their approval for the use of physical discipline (50.3 %, Table A1, Appendix A), in actual practice, only 7.7% applied the method (Table 4.6). Also, the calculated predicted probability that a household head will apply physical discipline is 0.03. This finding is at variance with the descriptive analysis of MICS 4 which indicated an increase in the usage of physical discipline by household heads (Ghana Statistical Service, 2011).

The low probability of using physical discipline methods compared to the other methods dispels the assumption about the traditional concept that physical discipline is the main disciplinary practice used in Ghanaian homes. In addition to showing that Ghanaian households, in practice, use a variety of discipline techniques, this finding also shows a possible change in attitude of Ghanaian parents' child disciplinary practices. The finding could also mean that household heads responses reflect desirability bias rather than actual behaviours. If so, this shift could reflect changes in parenting practices (an evolution of less violent approaches to discipline) around discipline but not cultural norms. By choosing more non-physical discipline alternatives it appears that Ghanaian parents are moving away from a commonly held practice of physical discipline toward more non-physical forms. This is noteworthy given that previous researches that did not consider clustering indicated high use of physical methods.

Table 4.7 shows the parameter estimates and standard errors from the null model (unconditional). Table 4.7 also shows four intercepts of equations of the different types of disciplinary methods simultaneously estimated. The equations represent the log odds of a household applying a specific discipline method. The category "All methods (random use of all methods or no preferred method") was used as the reference category. The negative intercept estimate (-2.4622) for choice of physical disciplinary methods category suggests that, the log-odds of households using physical methods in a typical region is significantly lower (Pr > |t| = 0.0006) than that of using the reference category.

Table 4.7 further shows that the estimates for the intercepts of a household applying non-physical discipline methods is -0.8921 and that for applying physical discipline methods, is -2.4622. These values imply that, for households in a typical region, the log-odds of using non-physical discipline methods, is lower than the log-odds of randomly using any disciplinary method (Reference category). Similarly, the log-odds of using physical discipline is lesser than that of randomly using any disciplinary method (Reference category). Furthermore, the log-odds of a household using psychological aggression discipline methods for correction is about 0.6 times lower (-0.5588) than randomly using any discipline method (Reference category).

4.5.3 Type of discipline methods applied across the regions

Table 4.8 presents information on the likelihoods of using the different discipline methods by households between the regions (our fifth research question). This is achieved by estimating the model's covariance parameters. The results show an intercept estimate of 0.2562 and standard error of 0.1216 with a p-value of 0.0175 for the application of non-physical and psychological methods. These figures mean there is significant variation between the regions in the log-odds of households using non-physical and psychological methods relative to the reference category (random methods). In terms of the likelihood of choosing to apply non-physical methods, the results (intercept = 0.2319, standard error = 0.2000, p-value = 0.0189) indicate significant variation in usage between the regions. Similarly, the estimates for physical discipline (intercept = 0.4193, standard error =0.1995, p-value =0.0178) and psychological aggression discipline (intercept =0.3620, std error =0.1691, p-value = 0.0161) indicate significant variation in their usage between the regions. In all these cases, the z tests suggest significant variation in the use of the methods relative to the reference between the regions.

Table 4.9: Covariance parameter estimates of the multilevel multinomial unconditional model

Covariance Parameter Estimates						
Cov Parm	Subject	Group	Estimate	Standard Error	Z Value	Pr > Z
Intercept	Region	Non-physical and Psychological methods	0.2562	0.1216	2.11	0.0175
Intercept	Region	Non-physical methods	0.2319	0.1117	2.08	0.0189
Intercept	Region	Physical methods	0.4193	0.1995	2.10	0.0178
Intercept	Region	Psychological methods	0.3620	0.1691	2.14	0.0161

Although not thoroughly studied, the region in which parents dwell seems to influence the likelihood that they use particular discipline methods. This suggests that the region in which families dwell constitutes some context that either supports or discourages the application of specific discipline methods. What is not clear however, is whether these effects could be attributed to cultural or social economic differences across the regions.

4.5.6 The conditional multinomial logit model results

The sixth research question sort to find out whether the study's covariates had any influence on the choice of the different categories of discipline methods. The conditional model analysis show the model results when the covariates were introduced in the model analysis. That is, whether the type of or category of disciplinary method used in households is associated with the covariates. At the level-1 stage of the analysis, 3 pairs of logits were modelled as a function of the covariates. At the level-2 stage, each intercept was modelled as a function of region. The intercepts were then varied randomly while examining changes in their objective function values. The objective function values are the -2 log likelihood (-2LL) value of the model in question.

Table 4.9 shows the estimated -2LL values of models M1 through to model M4. The model with the smallest -2LL value is considered the best fit for the data. Table 4.9 shows that, model M4 had the smallest -2LL value and therefore was applied to answer the research questions.

Table 4.10: Fit Statistics for the Conditional multinomial outcome model

Model	Effects Added	-2 log L
M1	(No predictors)	24,003.29
M2	M1 plus 3 predictors (sex of head, age-group of head, number of children	23,925.35
M3	M2 plus 2 predictors (religion of head , Education of head	22,835.80
M4	M3 plus 2 predictors (Wealth index ethnicity of head) ie (all predictors in model)	22,777.72

Covariate effects on the multinomial logit outcome

Table B4 in Appendix B shows the estimated parameters (model M4) of significant level-1 covariates on the types of different discipline methods under discussion. The Table shows that, generally, the categorical covariates, *number of children aged 2-14, religion of head of household and education level of household head had* significant influence on a household's likelihood to apply the different discipline methods (non-physical and psychological aggression, non-physical, physical and psychological aggression methods) relative to the reference category (use of random methods). Specifically, for the categorical covariate, *number of children aged 2-14,* households with 2 children (coefficient 0.2540, $p = 0.0147$) had significant influence on the use of non-physical and psychological discipline methods. This means that, households with 2 children aged 2-14 are more likely to use non-physical and psychological discipline methods than households in the reference category (3+ children).

The results further show that the category Moslem head of household, for *religion of head*, has significant and positive relationship with the different discipline methods (non-physical methods =0.6047, p = 0.0059, physical methods = 0.6687, p = 0.0132, psychological = 0.4937, p = 0.0238). This means, Moslem household heads are more likely to use the different discipline methods than household heads with 'other' as religious affiliation. For the covariate, *educational level* of *head,* the category, Middle/JSS has a coefficient estimate of 0.5585, and a significant p value of 0.0008, for use of non-physical and psychological discipline methods. Again, the category 'none' (no education), is highly significant (p < 0.0001) across all the different discipline methods. The 'primary' category of educational *level of head* had coefficient estimate of 0.4528 and p value of 0.0176 for non-physical and psychological aggression discipline methods. These figures mean household heads in these categories are more likely to use the different discipline methods than the household heads in the reference category (Secondary plus category).

In summary, predicting the odds that a household will use non physical and psychological aggression methods (relative to the reference category), that household should have 2 children or the household head should have no education.

The profile of households who are likely to use physical discipline and psychological aggression, are households where the head has no education (none) and whose religious affiliation is Moslem. These results are the net effect of all the other covariates.

Numerous studies show that socio-economically advantaged and disadvantaged parents differ on average in the quality and quantity of their interactions with children. The finding from this study suggests that, the number of children aged 2–14 in a household has a great influence on the type of discipline method a household is likely to use. This socioeconomic parenting gap could be the result of many factors including economic hardship, emotional stress, and lower access to information and services that often accompany low income. For example, Pinderhughes et al. (2000), found significant association between lower socio-economic status of families and usage of harsher discipline methods.

Discipline strategies differ by family background such that parents with lower education and lower income levels tend to use physical discipline more often, and use non-physical forms of discipline less often, than parents of higher socio-economic status. Gershoff (2002), suggests an additional explanation for these findings: "mothers with low education may more often believe that physical discipline prepares children for the larger world that will expect them to obey and not resist". Straus (2001) and Grogan-Kaylor and Otis (2007) concluded that "These mothers may not be aware of its negative effects on children's behaviour and their psychological and social well-being such as showing higher rates of criminal activity, perpetration of partner assault, abuse against one's own children and depression"

"The use of physical discipline on children has been supported by the religious affiliations and beliefs of many" (Kuczynski & Hildebrandt, 1997). This observation is common among, conservative Protestants with conformist religious beliefs and linked with more frequent use of physical discipline (Xu et al., 2000). The finding of this research agrees with the above studies as it indicates that Moslem household heads have a more likelihood of using physical discipline methods than their counterpart household heads in 'other' religion category (reference category). Religion of the household head in this instance was not indexed by household head's attendance at religious services but by merely indicating their religious beliefs. Further research is needed to assess the level to which religious affiliation and religiosity either separately or together predict parents' use of physical discipline methods.

Parents' ethnicity is considered as a tradition based on nationality, language, and culture. A number of researches have been conducted to determine if ethnicity is a possible determinant of the use of physical discipline (Betancourt & Lopez, 1993) or a relationship exists between ethnicity and disciplinary environment (Marshall, 2005). The results of these studies however, are "inconclusive and largely contradictory" (Smith et al., 2005). Some studies, however, have found parents use of physical discipline methods in certain ethnic groups more often than in other ethnic groups (Loeber et al., 2000). Where such variances have been identified, the effects are very small. In terms of actual use of discipline methods, the findings of this study indicate ethnicity has no significant relationship with use of physical discipline methods thus, supporting earlier findings.

The effect of parents' ethnicity on the type of discipline method likely to be used at home, is a tricky one as variations in the use of the methods have been observed to exist between ethnic groups and within ethnic groups. Parke and Buriel (1998), found that this variation is the result of the influenced of varying income levels within and between the groups.

4.5.7 Effects of significant level-1 covariates on the different types of discipline methods

This section discusses the associations between important influential covariates and the response variables using hypothesis tests. It answers the seventh research question on whether specific research covariates influence the usage or application of specific discipline methods differently.

4.5.8 Use of non-physical and psychological aggression discipline methods

Table B4 in Appendix B shows that, the point estimates and *p* values of two categorical covariate; educational level of household head and number of children aged 2-14 in a household show they are the most influential covariates in the determination of using non-physical and psychological discipline methods.

For the categorical covariate (educational level of head) household heads in the category 'none' (3.4043, $p < 0.0001$) are the most likely to use non-physical and psychological aggression discipline methods in contrast to the reference category (Secondary +). Table B4A further shows that the odds ratio for household heads (OR = 30.093, 95% CI: 19.492, 46.459) in this education category (none), is about 30 times more likely to use the method compared to their counterparts in the reference category (Secondary +). The second categorical predictor that influences the use of non-physical and psychological aggression discipline methods discipline method is the number of children aged 2-14 in a household. For this covariate, the category of 2 children in household, has a point estimate of 0.2540 and a *p* value of 0.0147. This means that households with 2 children are more likely to use non-physical and psychological aggression discipline methods compared to the reference category (3+ children). In terms of odd ratios (OR = 1.289, 95% CI: 1.051, 1.581), such households are about 1.3 times more likely to use these methods as a way of disciplining children compared to households within the reference category.

4.5.9 Use of non-physical discipline methods

In terms of applying non-physical discipline methods, two categories of the covariate educational level (None and Middle/JSS) and one category of the covariate religion (Moslem) have significant influence on household heads to use non-physical discipline methods. Estimated point estimates and *p* values of two levels (None and Middle/JSS) of the covariate educational level of household head namely the category 'none', shows a positive significant relationship across all the categories of discipline methods while household heads in the category Middle/JSS (0.5585, $p = 0.0008$) were more likely to use non-physical discipline methods than the reference category (Secondary+). The coefficient and *p* values of the category 'Moslem' (0.6047, $p = 0.0059$) of the covariate religion, is significant for non-physical discipline methods. This means Moslem household heads are more likely to use this method of discipline than the reference (other). In terms of odd ratios, a Moslem head is 1.8 times more likely (OR= 1.831, 95%, CI: 1.190, 2.815) to use this method than the counterpart head (reference category).

4.5.10 Use of physical discipline methods

With respect to using physical discipline methods, the results show the covariates educational level of household head and religion of household head are significant. Household heads in the 'none' category (with estimated coefficient, 3.9732 and $p < 0.0001$,) of the covariate educational level of head are the most likely to use this method of discipline. Meaning household heads with no education are more likely to use physical discipline method than the reference category (household heads with Secondary and more education). Again, Moslem household heads are more likely to use physical discipline methods than household heads in the 'other' category (reference category). Specifically, about 2 times more likely (OR= 1.952, 95%, CI: 1.150, 3.312) than the reference category.

4.5.11 Use of psychological aggression discipline methods

The application of psychological aggression discipline methods is significantly influenced by the education level of household head. Household heads in the category of 'none' are more likely to use any of the different types of discipline methods compared to the reference category. Again, household heads in 'Middle/JSS' education level are more likely to use non-physical discipline methods (0.5585, $p = 0.0008$) relative to the reference. The coefficient for the category Middle/JSS has a negative value (-1.050, $p < 0.0001$) meaning that a parent in this education level category is less likely to use psychological aggression discipline methods compared to parents in the reference category.

Finally, with the exception of the covariates number of children in household, religion and educational level of household head, all the other covariates appear to have no effect on the response variable.

4.5.12 Application of discipline methods across regions

The covariance parameter (cov parm) estimates for regions, indicate what variation, exists between the regions in the usage of the different discipline methods. That is, there is variability between the regions in terms of the likelihood of using the different types of discipline methods under discussion. Table 4.10 presents estimates for four intercepts; non-physical and psychological aggression discipline methods (0.0312, $p < 0.0569$), non-physical discipline methods (0.0299, $p < 0.0837$), physical discipline methods (0.0479, $p < 0.1045$) and psychological aggression discipline methods (0.0459, $p < 0.0510$).

Table 4.11: Multinomial Outcome Model: Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate	Standard Error	Z Value	Pr > Z
Intercept	Region	Non-physical and Psychological methods	0.0312	0.0197	1.58	0.0568
Intercept	Region	Non-physical methods	0.0299	0.0217	1.38	0.0837
Intercept	Region	Physical methods	0.0479	0.0382	1.26	0.1045
Intercept	Region	Psychological methods	0.0459	0.0281	1.64	0.0510

The estimates indicate that with the exception of use of psychological aggression discipline methods, which significantly varied between the regions, usage of all the other discipline types do not show any significant variation between the regions.

4.5.13 Regional effects on discipline methods

Table B5 in Appendix B shows level-2 (region) effects on each type of discipline method. It addresses the eighth research question on specific regional effects on the application of the different discipline methods. In determining the level-2 (region) effects, the household background variables were controlled and only the region effect was considered.

4.5.14 Regional effect on use of non-physical and psychological aggression discipline methods

Table B5 in Appendix B shows that when household background variables are controlled, the point estimates associated with households in Brong Ahafo (0.2829), followed by Ashanti (0.0.0875) and Western (0.0864) regions indicate higher log odds of applying non-physical and psychological aggression methods relative to use of the reference category (random us of all methods) than households in the other regions. The region with the lowest log-odds of households using non-physical and psychological discipline methods, relative to the reference category, than all the other regions is the Central region (-0.2969, p = 0.0611).

4.5.15 Regional effect on use of non-physical discipline methods

Table B5 in Appendix B again shows that between the regions, the point estimates of Upper West (0.2698, $p < 0.0484$), Western (0.0939, $p < 0.0368$) and Central (0.0633, $p < 0.5463$) regions are the highest; suggesting households in these regions have higher log-odds of using non-physical discipline methods, relative to the reference category, than the households in the other regions . The point estimate for Upper West region, and Western regions are significant while the estimate associated with households in the Central region is not significant. The Volta region (-0.1750, $p < 0.2020$) has the smallest point estimate (relative to the reference category), indicating the region with the lowest log-odds of households using non-physical discipline methods. In this instance, the point estimate associated with the Volta region failed to achieve statistical significance.

4.5.16 Regional effect on use of physical discipline methods

Table B5 in Appendix B also shows the Eastern region has the highest point estimate of 0.2457 and a p-value of 0.1773, implying households in the region have the highest log-odds of applying physical discipline methods (relative to the reference category) than the other regions This is followed by households in the Volta (0.2348, $p = 0.2135$) and North (0.0950, $p = 0.4935$) regions respectively. The region with the lowest log-odds of households applying physical discipline methods (relative to the reference category) is the Central (-0.3207, $p = 0.0667$).

4.5.17 Regional effect on Use of psychological aggression methods

In terms of applying psychological aggression methods, point estimates associated with households in Upper East (0.3038, $p < 0.0226$), followed by Ashanti (0.2276, $p < 0.1276$) and Northern (0.1094, $p < 0.3656$) regions are the highest. This suggest that, these regions have households with higher log-odds of using this method of discipline (relative to the reference category) than households in the other regions. The Greater Accra region with the point estimate of -0.2848 and a significant p value of 0.0467 is the region with households with the lowest log odds of applying psychological aggression discipline methods (relative to the reference category) than the other regions.

4.5.18 Multinomial logit model summary results

The results show that the different discipline methods have different probabilities of application. These probabilities vary between the different regions. Overall, the most likely discipline method that a household would use (relative to the reference method = random use of all methods) is a combination of non-physical and psychological aggression discipline methods. The log odds of households using physical discipline was significantly lower than using the reference category (random use of all methods). In all cases, the tests suggest significant variations between the regions in the log odds of using the different discipline methods relative to the reference method.

Significant associations between the study variables and the response variable were identified. At the household level, three categorical covariates (number of children aged 2-14, religion of household head and educational level of household head) indicated significant relationship with the response variable.

For the categorical variable, number of children aged 2-14 in household, households in the category of 2 children, were more likely to apply physical discipline methods than households with 3 or more children. For the categorical variable religion of household head, Moslem household heads were more likely to use non-physical and psychological discipline methods than household heads in the category ‘other’ (reference). For the categorical variable educational level of household head, households’ heads in the Middle/JSS level of education were more likely to use non physical and psychological aggression discipline methods (relative to the reference – secondary or more education). Again, household heads in the category of no education (none) were more likely to use all the different discipline methods relative to the reference category.

Holding the household background variables constant, the regional effects were as follows: Households in Brong Ahafo region had the highest log odds of applying non-physical and psychological aggression discipline methods than households in all the other regions. Households in Upper West region had the highest log odds of using non-physical discipline methods. Households in Eastern region had the highest log odds of using physical discipline methods, with households in the Upper East region having the highest log odds of applying psychological aggression discipline methods.

4.6 Model Evaluation and diagnostics (Binary logit Model)

This section examined the accuracy of the fitted binary logit model by assessing the ability of the model to adequately describe the variations in the response variables and by determining how close values predicted by the model are to the observed values. To achieve this, a 10-fold cross validation was conducted, followed by the construction of a classification table and ROC curves.

The final binary logit model for Objective 1 (Table 4.5 model M3) can be expressed as follows:

$$Y_{ij} = B_0 + \sum_{j=1}^{k-1} B_j E_{ij} + \varepsilon_{ij}$$

Where

Y_{ij}	The score on the response variable Y for subject i in group j.
B_0	The intercept that represents the grand mean of the response variable for all groups
k	The number of categories of the predictor variable
B_j	The regression coefficient associated with the jth group, represented by the difference between the mean of the group coded 1 on the corresponding dummy variable and the grand mean of all groups. In other words, it represents the effect of being in the jth group.

E_{ij} The numerical value assigned to subject i in the jth group

ε_{ij} The error associated with the ith subject in the jth group

4.6.1 Ten-fold cross validation results of the Binary logit model

Table 4.11 shows the estimated coefficients, the overall significance and partial significance of the variables included in the model when a 10-fold cross validation was conducted. The coefficients of the variables were estimated using the maximum likelihood estimation (MLE) method. The (MLE) method involves choosing values for the coefficients to maximize the likelihood (or probability) that the model will predict the same choices made by the observed responses. That is, the estimated coefficients are the values of the covariates which when plugged in the model will predict the same responses made by the observed responses of the households.

Table 4.12: 10-fold cross validation and parameter estimates of variables

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9034	0.7823	1.3335	0.2482
Head_Ethnicity	1	-0.0231	0.0236	0.9592	0.3274
Head_AgeGroup	1	-0.0166	0.3033	0.0030	0.9563
Num_Children	1	0.7687	0.3000	6.5643	0.0104
HHold_Windex	1	0.5756	0.1917	9.0183	0.0027
Head_Ethn*Head_Ethni	1	0.000180	0.000081	4.9310	0.0264
Head_Ethn*Head_AgeGr	1	0.00410	0.00453	0.8177	0.3658
Head_Ethn*Num_Childr	1	-0.00788	0.00557	2.0020	0.1571
Head_Ethn*HHold_Wind	1	-0.00103	0.00461	0.0503	0.8226
Head_AgeG*Head_AgeGr	1	-0.00912	0.0426	0.0457	0.8307
Head_AgeG*Num_Childr	1	0.0241	0.0442	0.2967	0.5860

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Head_AgeG*HHold_Wind	1	-0.0297	0.0329	0.8115	0.3677
Num_Child*Num_Childr	1	-0.1189	0.0514	5.3450	0.0208
Num_Child*HHold_Wind	1	-0.0602	0.0379	2.5191	0.1125
HHold_Win*HHold_Wind	1	-0.0839	0.0155	29.4411	<.0001
Head*Head*Num_*HHold	1	0.000407	0.000692	0.3463	0.5562

In fitting the model for the cross validation, predicted probabilities for each observation was generated. To assess the accuracy of this prediction, i.e., the ability of the model to distinguish correctly the two classes of outcomes, a 2 by 2 table was constructed with the rule that if a predicted probability for an observation was more than 0.50 it was classified as a positive test for the response variable, otherwise it was classified as a negative test.

Table 4.12 shows the constructed 2 by 2 table with 2,652 true positives and 1,770 true negatives. In terms of prediction error, 2,283 observations were classified as 'yes' when in fact they were 'no'. Similarly, the model classified 1,451 observations as 'no' when in fact they were 'yes'. This gave a sensitivity of 64.64% (2,652 / 4,103) to indicate the measure of the percentage of positive households which were classified as positive. In other words, the model provides a good classifier for the decisions of the households and can assist in estimating the correct probability of a household to fall in the two categories.

Table 4.13: Classification of Binary responses (Test results)

Support Physical Discipline	No	Yes	Total
Yes	1,451 35.36%	2,652 64.64%	4,103
No	1,770 43.67%	2,283 56.33%	4,053
Total	3,221	4,935	8,156

Table 4.12 again shows a specificity of 43.6% (1,770 / 4,053) which measures the percentage of negative households which were classified as negative. This means the model has the ability to identify households who do not approve the use of physical discipline as proper for good child upbringing.

A more complete description of classification accuracy is given by the area under the ROC curve where it is possible to judge the effects of sampling variables on the estimate of the AUC.

The ROC curve, is a plot of the probability of correctly classifying a positive subject against the probability of incorrectly classifying a negative subject for the entire range of possible cut-off points. The area under the curve (AUC) range from 0 to 1 and provides a measure of the models ability to discriminate. The larger the area under the ROC curve the more the model discriminates.

Table 4.13 displays measures of association between predicted probabilities and observed responses, which include a breakdown of the number of pairs with different responses. Table 4.13 indicates that out of the 16,629,459 pairs, 55.2 % were concordant meaning in a randomly selected pair of households, the one with a higher probability actually agrees to a 'yes' and the one with a lower probability also agrees to a 'yes' and 1.9% were tied. Because the outcome is binary, the 'c' statistic represents the area under the ROC curve. The ROC curve represents a plot of the values of sensitivity against 1 minus specificity as the value of the cut-off point 'c' is increased from 0 to 1. That is, it shows the tradeoff between the sensitivity and specificity as the decision threshold varies.

The area under the ROC curve is a measure of quality of a probabilistic classifier. A random classifier has an area of 0.5, while a perfect classifier has an area of 1. In other words, the value 0.562 (Table 4.13) means 56.2% of the area under the ROC curve. The Gamma statistic (Table 4.13) indicates a value of 0.126 meaning that 12.6% fewer errors were made during the prediction. Table 4.13 also reports Somers' D statistic which is related to the concordance by means of D = 2*(c-0.5) and gives a value of 0.123. Somers' D is the rescaled version of concordance that takes values between -1 and +1, like the usual correlation coefficient instead of 0 and 1.

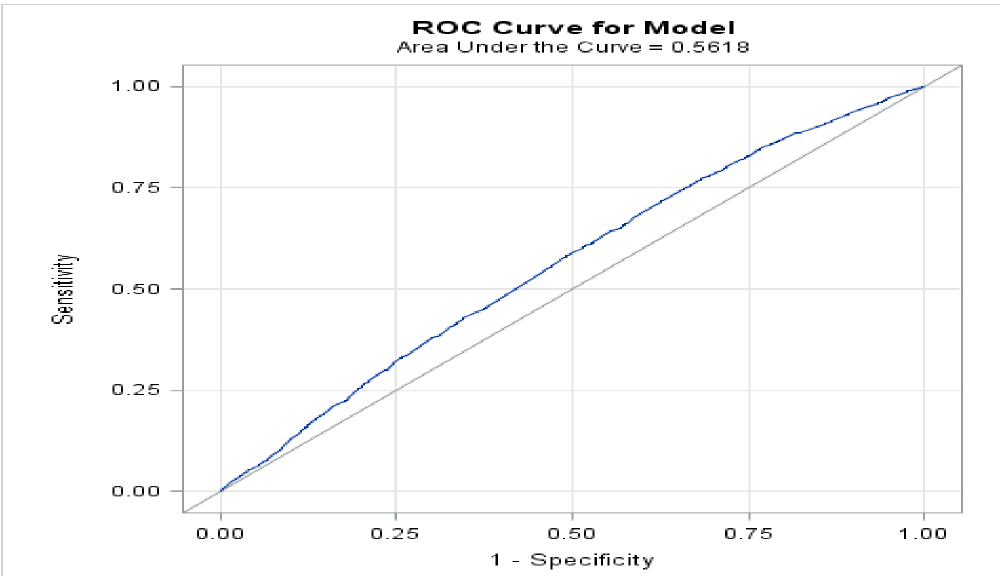
Table 4.14: Association of Predicted Probabilities and Observed Responses

Percent Concordant	55.2	Somers' D	0.123
Percent Discordant	42.9	Gamma	0.126
Percent Tied	1.9	Tau-a	0.062
Pairs	16,629,459	c	0.562

Figure 4.1 shows the area under curve (AUC) when the binary logit model with significant covariates was applied to the training dataset at the 0.05 significant level using cross validation. Figure 4.1 shows an AUC of 0.5618 which translates to 56.18%

suggesting a better classification than the uninformative model with an expected area of 0.5 (50%). The area under the uninformative model is the area of the model containing only the intercept and no covariates when fitted. This area is 0.5 (the area under the 45 degree line).

Figure 4.1: AUC for Binary logit model with significant covariates fitted on the training dataset using 10-fold cross validation

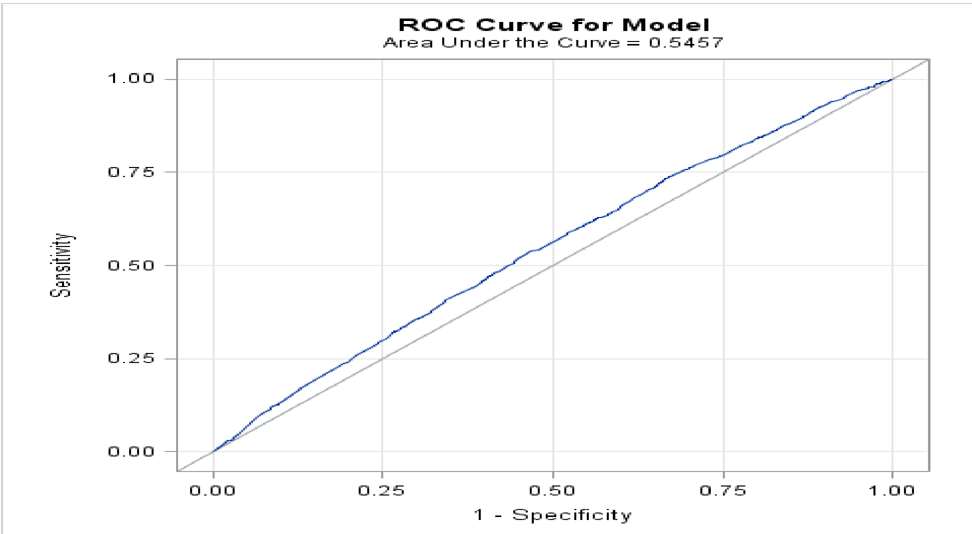


To determine whether the observed difference in AUCs (ie between the developed binary logit model and the uninformative model) is significant and not by chance, an ROC contrast test was carried out to further analyze this difference. The ROC contrast test indicated the two AUCs differed by 6.18% (56.18% - 50.00%) with a standard error of 0.01% (0.0091). A significance test for this difference showed a confidence interval of 4.02% to 8.34%, with a *p*-value of < 0.0001. The confidence interval did not include zero, indicating that the two predictions are statistically distinguishable and did not happened by chance.

4.6.3 Comparing the ROC curves when the model is applied to the validation dataset

Figure 4.2 shows the AUC of the binary logit model after applying the model to the validation dataset. The AUC in this instance was 0.5457 which is lower compared to when the model was applied to the training dataset. This drop, (from 0.5618 to 0.5457) with a significant $p < 0.0026$ indicates the model will discriminate well on a new dataset.

Figure 4.2: ROC analysis of validation data using 10-fold cross validation



In conclusion, our binary logit model can be said to have met the criteria to discriminate well on a new dataset because it indicated a lower level of performance when shifting from fitting the model to the training dataset to the validation sample. Thus, this slight reduction or drop in the magnitude of the performance measure is expected. However, if the drop in value of the measure was too large, the model would then be considered as not being able to validate outside the fitting sample.

4.6.4 Comparing the Generalized linear mixed model (GLMM) and the Logistic model

A GLMM combines the characteristics of generalized linear models and mixed models where the model considers both fixed and random predictor variables. Figure 3 shows four charts, A, B, C and D of ROC curves and areas under the curves for fitted logistic models for the dataset when random effects (GLMM) are considered and when random effects are not considered, in one plot. The AUCs of the ROCs were used to summarize the discrimination ability (rate of successful classification) of the two models.

Charts A, B, C, and D show that the AUC for a model which considers random effects because of the hierarchical nature of the dataset has a higher discrimination ability than the model that did not factor in random effects. The charts show that the AUC for the GLMM is better than that for the logistic model in all selected proportions of the training and test/validation data. Chart A, specifically, shows the ROC curves using 60% of the data as training data and 40% as the test/validation data. The chart, shows the AUC for the GLMM is bigger than the AUC for the logistic model. Chart B (70% training data and 30% test data), Chart C (80% training data and 20% test data) and Chart D (90% training data and 10% test data), similarly show the ROC curves for the GLMM have better classification than that of the logistics model.

In conclusion, when there is non-independence in the data, such as arises from a hierarchical structured data and the outcome variable has some distribution that does not result in normally distributed errors, it is best to use GLMM which takes into consideration, the differences between groups. However, if one assumes the variables are independent when they are not, and tries to fit non-normal data with a normal distribution, estimated *p*-values and standard errors will be incorrect and obtained statistics will be biased.

Figure 4.3: ROC curves and areas under the curves for fitted logistic models when random effects (GLMM) are applied and when random effects are not applied.

Chart A: 60% -Training and 40%-Test Chart B: 70% -Training and 30% -Test

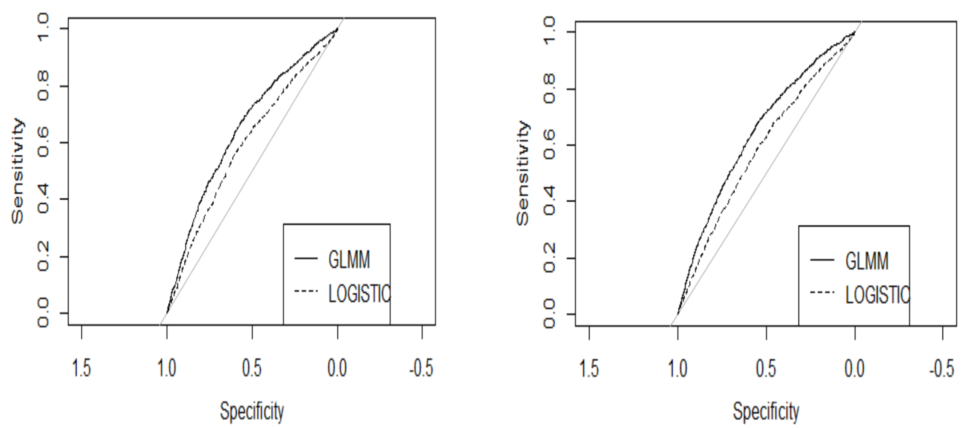
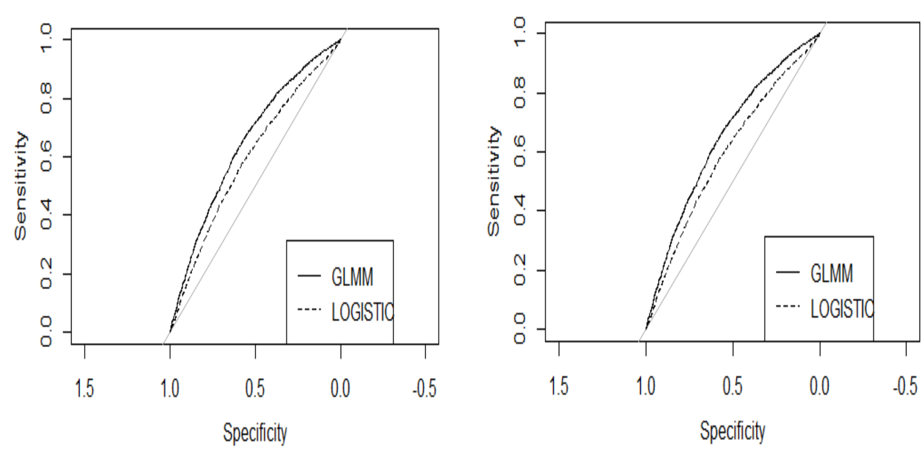


Chart C**: 80%-Training and 20%-Test** Chart D**: 90% -Training and 10% - test**



4.6.5 Simulation Results for the Binary logit model

Regression models describe how the response variable is related to the explanatory or predictor variables. For logit models, this relationship become more complicated as it now involves a logit transformation that relates the mean response to the explanatory variables. Because the explanatory variables are categorical, the generalized linear mix model procedure was applied in the model simulation.

The simulation involved generating data of different sample sizes using the developed binary logit model and performing regression analysis to validate the correctness of the simulated data to determine how robust and efficient the developed model is.

The simulation steps involved; simulating different sizes of samples, determining the characteristics of the explanatory variables, simulating the binary logit model with explanatory variables and performing a logistic regression procedure. Finally, the generated parameter values from the simulation were compared with the true parameter values to determine how close they were to the true parameter values. Also, an assessment of whether a generated 95% confidence interval for the estimated parameters of the simulated samples contained the true parameter values.

Table 4.14 shows the maximum likelihood estimates of parameters of simulated sample sizes varying from 1000 to 5000 from the dataset using the binary logit model.

Table 4.15 Parameter estimates for 1000 to 5000 simulated observations with the binary logit model

Parameters	True parameter values	Simulated sample size N				
		N=1000	N=2000	N=3000	N=4000	N=5000
Intercept	-0.4121	-0.438	-0.1117	-0.3867	-0.6325	-0.2523
Age group of head	-1.0438	-1.4049	-1.2747	-1.0898	-0.9086	-1.2343
Ethnicity	0.6850	0.9393	0.3442	0.6837	0.6648	0.7589
No. of Children	-0.2124	-0.2232	-0.3224	-0.328	-0.1187	-0.4208
Wealth Index	-0.5165	-0.4927	-0.4526	-0.457	-0.2518	-0.4757

Figure 4.4 shows that there is little variation between the values of the estimated parameters and the true parameter values of the covariates. Figure 4.5 also shows that as the simulation sample size increased the standard error in estimating the parameters from the samples kept reducing and thus indicating the model's ability to provide improved estimates of the true parameter values.

Figure 4.4: Variation in estimated parameter values compared to true parameter values with increasing simulated sample size.

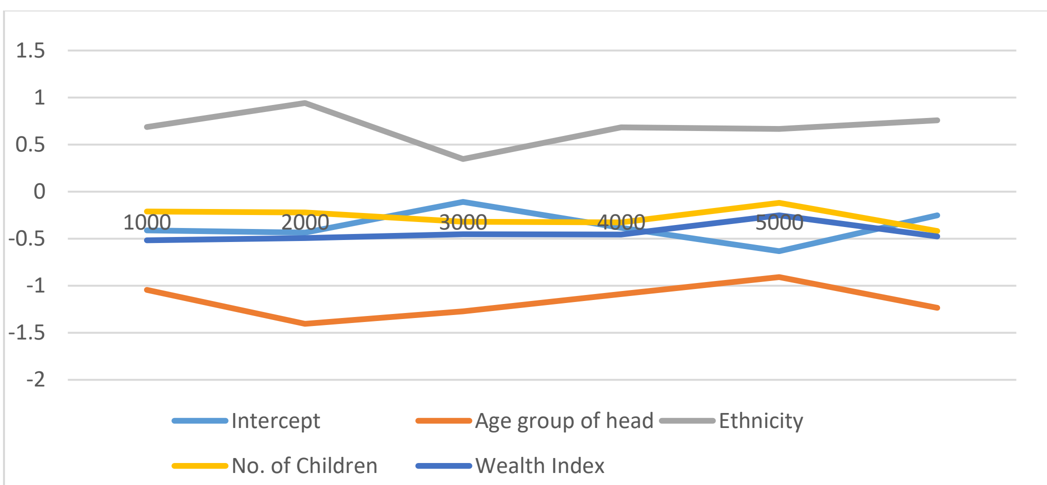
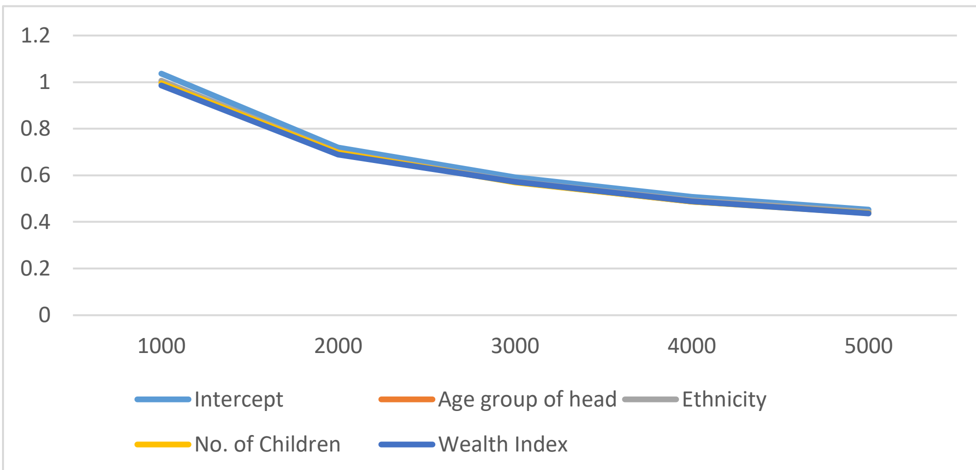


Figure 4.5 also shows the range (distance) between 95% CI limits for estimated parameter values with increasing simulated sample sizes. The range of the 95% CI for each estimated parameter included the true parameter values and the intervals reduced with increasing simulated sample sizes leading to high precision in the estimate of the true parameters.

Figure 4.5: 95% Confidence Interval limits for estimated parameter values with increasing simulated sample sizes



4.7 Model Evaluation and diagnostics (Multinomial Logit Model)

Two approaches were used in this evaluation; Creation of a classification table and applying the Bootstrap method.

4.7.1 Classification table (Confusion matrix) results

The ability of the multinomial logit model to classify correctly new data into the different discipline categories was assessed. The goal of the classification was to identify the category or class to which a new data will fall using the multinomial logit model. Table 4.15 shows the classification results (confusion matrix) of applying the final multinomial logit model with the significant covariates (number of children, religion of household and education level of household head). Table 4.15, show the cross classification of the observed and predicted response variables for the data used to fit the model. It also shows when the observed and predicted response levels agree and the percentage correct.

Table 4.15 also shows that the multinomial logit model could correctly classify households which used non-physical discipline methods with an accuracy of 31%, households who applied psychological aggression discipline methods with an accuracy of 58%, households who applied a combination of non-physical and psychological aggression discipline methods with an accuracy of 85% and households that randomly applied any discipline method with an accuracy of 65%. The model, however, could not classify households who applied physical discipline methods. The classification in this instance was zero percent.

The overall correct classification rate of the final multinomial logit model, however, was 56% (Table 4.15). The apparent low rate of correct classification, particularly in terms of physical discipline methods could be due to the fact that the number of elements (households) in that category was very small. This means that there are still other risk factors that have not been accounted for. Table B10 in Appendix B show that the initial classification results with no covariates in the model, had an overall percentage accuracy of 33.9%.

Table 4.16 Classification table for the multinomial model with significant covariates

Observed	Predicted						
	Non-physical methods	Psychological methods	Physical methods	Non physical and Psychological methods	All methods	Total	Percent Correct
Non-physical methods	578	331	0	906	132	1,877	31%
Psychological methods	179	1,238	0	670	136	2,153	58%
Physical methods	66	145	0	209	205	625	0%
Non physical and Psychological methods	176	262	0	2,204	43	2,602	85%
All methods	91	70	0	156	584	899	65%
Total	1,020	1976	0	4,065	1,095	8,156	56%

4.7.2 Bootstrap results

To determine whether the model is efficient, the bias in the estimates of the coefficients of the significant covariates using bootstrap with replacement and construction of a 95% confidence interval for the estimates were assessed. The bias is the difference between the true parameter value (reference) and the observed average of the estimates on the same characteristics in predicting the response variable.

After first determining the reference values (original parameter values) by computing the statistics for the original data, 20 bootstrap samples (with replacement) were obtained and the statistics of interest computed for each bootstrap sample. The union of these statistics constitutes the bootstrap distribution which approximates the sampling distribution of the statistics under the null hypothesis of bias and confidence interval.

Two approaches were used to analyse the results of the bootstrap estimates; by graphically using a histogram and numerically developing a confidence interval around the average of estimated coefficients of the significant covariates. Table B5A in Appendix B shows the bootstrap estimates of the significant parameters sampled from 20 repetitions and using 50% (4,078 observations) of the overall data. These results were used to determine if the estimates have any bias. Figures 4.6, 4.7 and 4.8, show histograms of estimated sampled parameters for significant covariates.

Figure 4.6: Histogram of Bootstrap resampling results of Educational level of household head

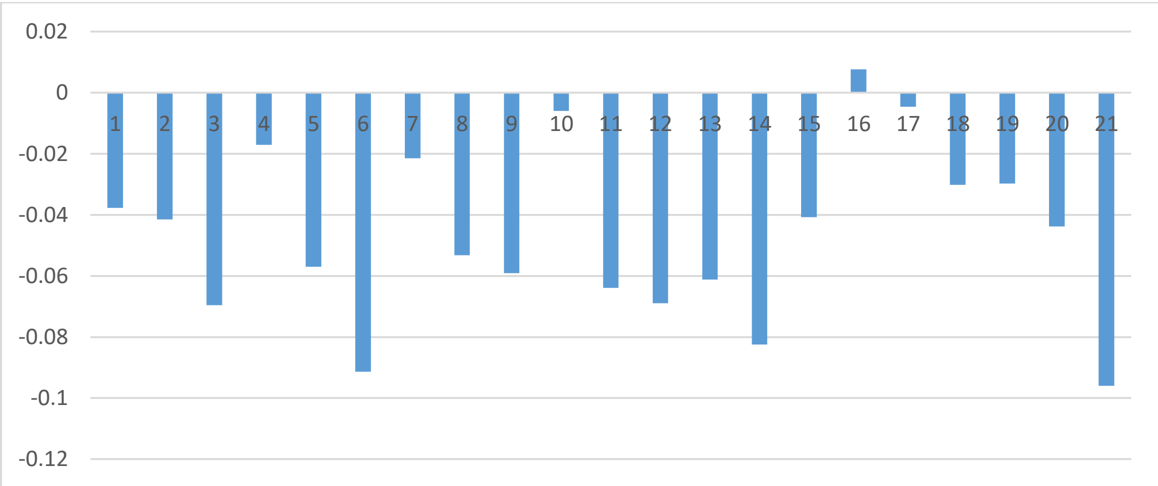


Figure 4.7: Historgram of Bootstrap resampling results of Number of children in household

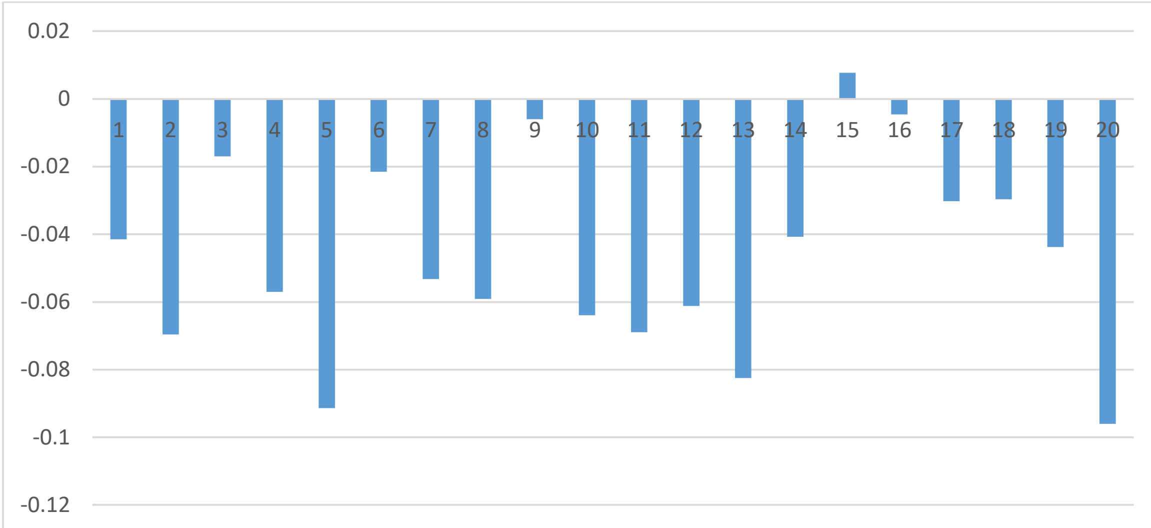


Figure 4.8:Historgram of Bootstrap resampling results of Religion of household head

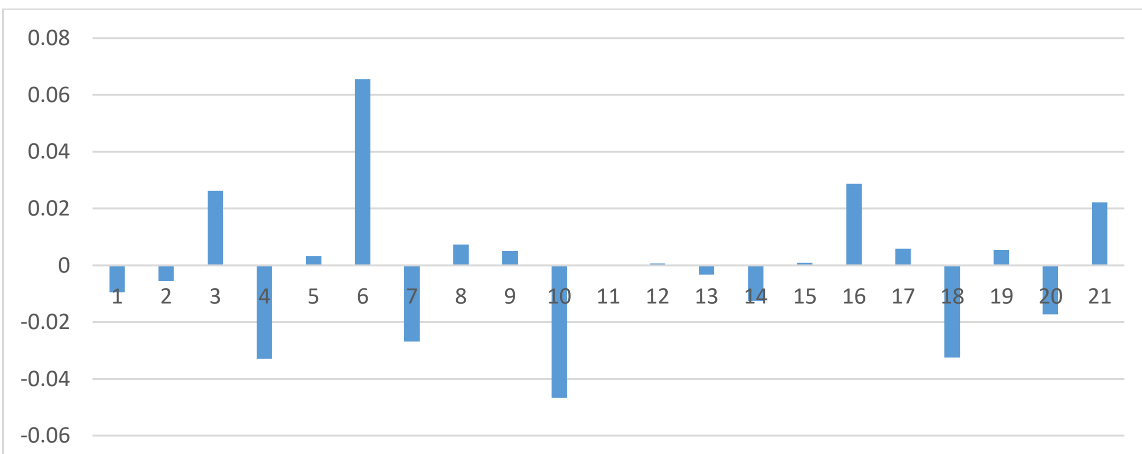
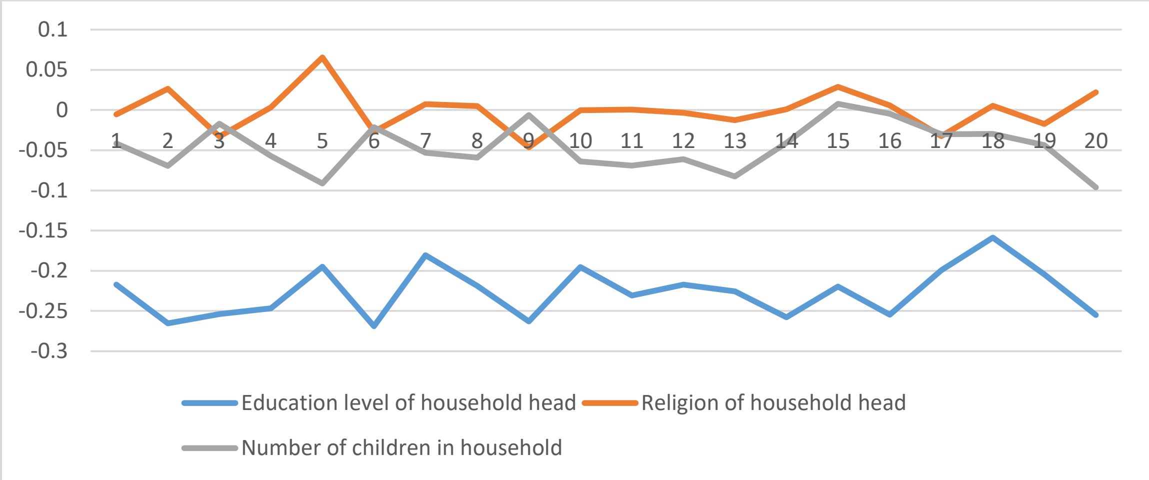


Figure 4.9: Estimated values of coefficients of significant covariates over 20 repetitions of bootstrap sampling with replacement



The histograms show that the results are spread out around the reference values (true values). In other words the bootstrap distribution appears to be normally distributed. This indicates that the bootstrap estimates would likely be similar to the reference parameter estimates, indicating absence of bias.

Figure 4.9 shows that the estimated values of the parameters from the 20 repeated bootstrap samples were not far from the true parameter values as the estimated values hovered around the true parameter values. Specifically, point estimates for the coefficient of the variable educational level of household head (-0.2261) fell in the interval (-0.2688 to -0.1587) and that for the variable religion of household head (-0.00948) fell in the interval (-0.0466 to 0.0656) and that for number of children in household (-0.0377) fell in the (-0.0914 to 0.0077). All these intervals contain the point estimates of the coefficients of the significant covariates.

To be sure that there was no bias with the bootstrap estimates, the precision of the estimated values were further checked by constructing 95% confidence intervals for the average estimated values of the coefficients of the significant covariates. The 95% confidence interval of a sample statistic is well approximated by the 2.5th and 97.5th percentiles of the sampling distribution of that statistic. If the confidence interval includes the reference value, then we confirm that there is evidence of no bias being present.

Table 4.16 shows the computations of the confidence intervals for the estimated values of the significant covariates at alpha level of 0.05. Since the confidence intervals contain the reference values, we conclude that there is no evidence of bias and the bootstrap estimates are approximates of the sampling distribution of the statistics under the null hypothesis.

Table 4.17: Confidence intervals for estimated values of significant covariates

	level of Education of household head	Religion of household head	No of Children in household
True parameter estimates (x)	**-0.2261**	**-0.0095**	**-0.0377**
Average of n =20 bootstrap samples estimated result ($\hat{x}$)	-0.2264	-0.0003	-0.0465
Bias = $\hat{x} - x$	-0.0003	0.0092	-0.0088
Standard deviation (s= 0.2207)	0.2207	0.2207	0.2207
Degrees of freedom (n-1 =19)	19	19	19
t value for t distribution with df =19 and alpha = 0.05	2.093	2.093	2.093
Upper confidence limit = $\hat{x} + t(\frac{s}{\sqrt{n}})$	-0.1231	0.0106	0.0568
lower confidence limit = $\hat{x} - t(\frac{s}{\sqrt{n}})$	-0.3297	-0.1036	-0.1498

4.8 Summary - Model Evaluation and Diagnostics

4.8.1 The binary logit model

Using a cut-off of 0.5 the binary logit model has a sensitivity of 64.6% and specificity of 43.6% and classification accuracy of 54.2%. This means the binary model is able to distinguish well, the households with different responses.

The AUC for the fitted binary logit model applied to the training dataset was 0.5618 with an ROC contrast test value of $p < 0.0001$ indicating the fitted model is better than the uninformative model. Also, the AUC for the fitted binary logit model applied to the validation data was 0.5457 indicating a slight drop in value after cross validation. However, the ROC contrast test with a $p < 0.0001$ value, indicated the obtained AUC was still significant despite the drop. Consequently, the model can be described as one that correctly discriminates on the original sample and on samples outside it.

Overall, the results indicate that the binary logit model has a discrimination property which is not specifically tied to the available sample, but could be extended to the whole population where the sample was selected from. In other words, the model is likely to discriminate well on data which is new and independent from the data used to fit the model. The Binary logit model, therefore, has a good fit and can adequately describe the relationship between household responses and predictors in the population.

4.8.2 The multinomial logit model

The overall correct classification rate of the final multinomial logit model, was 56%. The initial classification results with no covariates in the model, had an overall percentage accuracy of 33.9%. The apparent low rate of correct classification, particularly in terms of application of physical discipline methods could be due to the number of elements (households) in that category which was very small. This observation also means there are still other risk factors that have not been accounted for.

The bootstrap results of the multinomial logit model show no evidence of bias in the estimated parameters as they approximate the sampling distribution of the model's statistic under the null hypothesis. This means that the coefficients of the model are not dependent on any particular portion of the original data used to fit the model. This is a sign of internal stability of the model.

CHAPTER 5

CONCLUSION

MAJOR FINDINGS, RECOMMENDATION, FUTURE RESEARCH DIRECTION AND CONTRIBUTION TO KNOWLEDGE

5.1 Introduction

This chapter provides information on the major findings, recommendation, future research direction and contribution to knowledge derived from this study. The study's results proved that using a multilevel modelling approach to analyze the data was the best to achieve the goals of the study. The study concluded that most important risk factors associated child discipline practices in Ghanaian homes (physical discipline against children) were not only directly related to the characteristics of the household to which children belong, but also the environment of the household.

The bivariate analysis of Objectives 1 and 2 of the study, showed that not taking into consideration the hierarchical levels of the data, nearly every predictor variable in the study indicated a significant association with the response variables which a multilevel in reality was not the case.

5.2 Objective 1 Conclusions

The results of the binary logit model which assessed the approval or otherwise of physical discipline as appropriate for child upbringing, showed that the predicted probability of approval of physical discipline by a household in the overall sample was 0.48. This answered the first research question on the probability of a 'yes' response to approval of physical discipline by a household. For the second research question on whether there was variability in the responses across the groupings or levels of the data, the results indicated that there was significant variability in the log-odds of a 'yes' response between the regions but not at the location (urban/rural) level. The third research question was whether there was any relationship, between household characteristics and approval for the use of physical discipline. The results showed that being in the age group of 15-21 years was associated with a lower predicted log odds of approving the use of physical discipline methods relative to being in the reference age group (61+). Also, ethnic groups like Akan, Ga/Dangme, Gruma and Mole Dagbni were associated with having a higher predicted log odds of approving the use of physical discipline methods relative to other ethnic groups. Again, households with one child aged 2-14, were associated with a lower log odds of approving use of physical discipline relative to the reference group (two or more children). Furthermore, households in the richest quintile of the Wealth index have lower predicted log odds of approving the use of physical discipline relative to the reference quintile (second).

5.3 Objective 2 Conclusions

From the multinomial logit model which considered the type of discipline methods actually applied by households in the survey, the analyses showed that the different discipline methods have different probabilities of application between the regions. The most likely discipline method that a household would apply relative to using just any method (reference), was a combination of non-physical and psychological aggression methods. The log odds of using physical discipline methods was lower than randomly using any method (reference). The results further indicated that significant variation exist between the regions in the application of all the different types of discipline methods (research question five).

To answer the sixth research question on whether the choice of the different categories of discipline methods depended on the research predictors, the study's results indicated that at the household level, the choice of a discipline method was influenced by three categorical predictors: the educational level of the household head, the number of children aged 2-14 in a household and the religion of household head. These categorical predictors had statistical significant effects on household's application of all the different categories of discipline methods.

Specific levels within these categorical variables that had effect on the responses were: For number of children in households - households with two children aged 2-14 were more likely to apply physical discipline methods relative to the reference category (3+). For religion of household head - households with Moslem heads were more likely to apply non physical and psychological aggression discipline methods relative to the reference category (other). For educational level of household head - households' heads with Middle/JSS level of education were more likely to use non physical and psychological discipline methods relative to the reference category. Again, household heads in the category of no education (none) were more likely to use all the different discipline methods relative to the reference category (Secondary +).

At the regional level, the results indicated that, households in Brong Ahafo had the highest log odds of using a combination of non-physical and psychological aggression discipline methods compared to households in the other regions. The region where physical discipline was most used was the Eastern region, with psychological aggression discipline methods mostly applied by households in the Upper East region. The Upper West region was the region where non-physical discipline methods were mostly used.

Diagnostic and Evaluation Conclusions

The diagnostic and evaluation results of the developed models indicated the binary logit model has a good fit and can adequately describe relationships between household responses and predictors in the population. The multinomial logit model with the significant predictors had an overall correct classification rate of 56%. Without the predictors (null model) the classification rate was 33.9%.

Taken together, these conclusions inform and deepen our understanding of parenting behaviour in terms of child discipline. Given these outcomes, a lot more effort is needed by stakeholders to disabuse the minds of parents and caretakers that physical discipline is appropriate for good child upbringing especially with the vast body of evidence linking these practices to detrimental child outcomes. It is hoped that the empirical findings from this study will assist and support social policy-makers as they plan guiding principles for positive parenting and develop programmes to raise awareness among Ghanaian parents on healthy disciplinary methods.

This study has also provided an in-depth analysis of child discipline data using a multilevel approach which addresses the problem of dependencies between individual observations in survey research where the sample is not taken randomly but instead from cluster sampling from geographical areas. If one has access to individual-level data on both the response and its potential predictors, as in the case of the MICS 4 dataset, a preferable approach should be to analyze the responses using a multilevel approach to explore the effects of the individual and group level covariates on the beliefs and practices of households, while simultaneously allowing for effects of unobserved group characteristics.

The application of physical discipline methods in the name of correcting children and guiding them to behave appropriately in society, still remains a controversial topic in the domain of parent-child relationship (Holden, 2002). The debate on their usefulness is still on-going. However, increasing number of countries are now taking a stance against the use of aversive discipline methods and legally banning their use by its citizens (Save the Children, 2002). There is the school of thought which says that if an adult hitting an adult is considered wrong and legally indefensible, then an adult hitting a child should also be considered improper.

5.4 Disparities between households' approval of physical discipline method and actual application of method.

The results show disparities in the two main responses of the study. In the first response (binary logit model), the percentage of households who approved the use of physical discipline was 50.3% whereas the percentage of households who actually used the method, the second response (multinomial logit model), was 7.7%. Again, the type of covariates that influenced household's approval for use of physical discipline in the first response was different from the type of covariates that influenced the actual application of physical discipline method in the second response.

Research shows that approval for use of physical discipline are based on personal experiences like; 'I was spanked and I am OK', 'Spanking improves child behaviour' , 'Other forms of discipline not effective as spanking'. For such parents there is the belief that today's generation or society is somehow worse off than their generation. This is because to them social changes have diminished physical discipline and this has led to lack of discipline, no fear or respect for authority, disrespect and lazy behaviour.

Although, the approval rate for use of physical discipline in the study was high and varied across the groups, in practice only few parents actually applied the method. Could this be the consequence of changing social norms as a result of parental access to relevant information and thereby shifting attitudes regarding the use of physical discipline? This question and others need to be answered in further research to draw concrete conclusion on these disparities.

The overall results show that for households who approved the use of physical discipline, the determining predictors were: age group of household head, ethnicity of household head, number of children in a household aged 2-14 and wealth index of household. For households who actually applied physical discipline methods, among the key overriding variables that influenced choice of physical discipline methods, was the number of children in household.

The number of children in a household appears to be the most significant factor in a households' decision to approve or use physical discipline methods for correction. This observation could be related to the challenges and demands associated with being a parent in a household with large number of children, and the responsibility of taking care of them, especially if one also has to contend with the stresses of a low socio-economic status.

5.5 Recommendations

In the light of the findings of this study, I recommend multilevel modelling in child discipline studies as the context in which such data are collected tend to be hierarchical in nature. A multilevel modelling approach, will account for uncaptured higher level variation, and adjust estimated errors in the estimation of standard errors to eliminate overstating the significance of estimated statistical relationships and falsely uncover statistically significant relationships when they do not exist. a

Despite efforts to eliminate all forms of violent discipline, there still persist physical and psychological aggression towards children. Multi-disciplinary approach programs should be established for parents and for poor families as a part of any strategy that aims to reduce the level of violence against children through discipline. At the same time, programs to identify as early as possible individuals with characteristics predisposal to violence, for example, individuals in low socio-economic groups who are at high risk for use of violence should be developed. There should also be programs to systematically eliminate cultural norms and media influence that legitimize and glorify violence.

Strategies to further reduce reliance on physical discipline in lower social economic households could include encouraging the use of non-physical discipline methods such as time-out and positive reinforcement, as well as raising public awareness about its negative effects, and relative ineffectiveness.

Moving forward, besides encouraging non-violent discipline, it is important to better understand which type of households are less likely to use non-violent discipline, why they are less likely to use them, and what targeted intervention efforts would be most effective in promoting the use of non-violent discipline. In this wise, priority should be given to less-educated parents because they are more likely to choose physical discipline methods than educated parents. More attention should also be given to older parents as they tend to apply more physical discipline methods than younger parents.

Further research should be conducted to consider the inclusion of additional covariates like whether the selected child for the study is the biological child of the household head and gender of the selected child which are related to child discipline practices. Such additions could provide much in-depth information to achieve a model with higher classification rates and robustness.

Finally, the results of this study indicated that the risk factors associated with violence against children, in terms of child discipline, were related not only to the parent's circumstance but also the environment of the household in which the child dwells. This suggests the totality of the family system should be the main focus of any engagement with parents and any policy to end violence against children in the name of discipline.

In view of these findings further research is needed to determine the source of regional differences/variation in the application of non-physical and psychological aggression methods and why no differences exist in the use of physical discipline across the regions. Intervention programs aimed at discouraging the use of physical disciplinary methods may need to involve different approaches, including promoting parents' discussion of positive parenting, home visitation programs (Dubowitz et al., 2012). However, Global efforts should be made in understanding where risk may be distinguished from harm without imposing one set of cultural practices to contexts where they are not helpful. Also, efforts such as media-based interventions may be effective in shifting social norms and parental discipline practices i.e. using printed materials (Reich et al., 2012), offering information through public education campaigns (McKeown, 2006).

5.6 Contribution to research and knowledge

There exist limited in-depth analysis of child discipline data in Ghana. This study provides a framework for discussion on structured data and approaches to their analysis. I am not aware of any previous analysis on child discipline data that provide such details using multilevel modelling.

Despite the limitations indicated in section 1.6 this study's results contribute to the growing knowledge on child discipline practices which suggests that Ghanaian children may be subject to psychological and physical abuse within the context of parental discipline or socialization. Although further research is needed to establish the actual regional effects (i.e. what in the regions cause these influence) on child discipline practices and to enhance our understanding of the consequences of child abuse in Ghana, further information has been provided to address some of the risk factors identified in this study. For instance improving the wealth index and educational level of household heads and family planning to manage the number of children in households, would bring many benefits to the Ghanaian child in terms of health, growth and development.

This study also has implications for counselor educators and for professional counselors as it provides an understanding of parenting practices in Ghanaian homes and the types of disciplinary techniques used in their cultural context. Based on the results of the study, it is important that counselors take into consideration the current multiplicity of issues that influence how Ghanaian parents and caregivers discipline their children. If we are to change social norms about physical discipline we must better understand the disconnect between the empirical evidence which demonstrates the harms of physical discipline and the highly prevalent attitudes and beliefs that support its use. With better understanding of the roots of prevailing norms regarding physical discipline, we can develop improved public health prevention and health education strategies designed to counter false beliefs and highlight alternative discipline approaches. Therefore, it is pertinent that counselors frame child discipline not only within the family but in addition other contextual factors. Thus family development should be the core of any policy program to end physical discipline and violence against children in the Ghanaian society.

Other important contribution of this study is, it will inform social policy-makers and designers of child maltreatment preventive programs of the status of child discipline practices in Ghana and provide a useful basis for the development of appropriate programs on contemporary child development issues. This study will also facilitate the development of targeted policies and programs on the dangerous outcomes of using aversive discipline methods and assist in the exploration of effectiveness of awareness creation on positive discipline. These findings will again offer a basis for counselors, school and social welfare personnel to initiate reframing their assumptions about child discipline in the Ghanaian family. To assess whether there has been real changes in child discipline methods among Ghanaian parents, these findings are important additions to the literature about Ghanaian parents and their utilization of a variety of disciplinary methods.

Statistically, the findings go to confirm the need to consider the structure of the data in order to choose the appropriate statistical analysis approach. Stratification in sampling ensures appropriate sample representation on the stratification variables, but yields too small (negatively biased) estimates of the population variance. If these aspects of complex survey data are ignored, standard errors and point estimates become biased and thereby potentially lead to incorrect inferences being made. So, not considering the hierarchies or clusters in a dataset and naively analyzing the data is likely to lead to a misinterpretation of obtained estimates of parameters and not show the reality of what is on the ground and thereby lead to wrong intervention by policy makers.

The description of a location as an urban area, connotes a place with availability of certain amenities (social and economic) that affect the life of the people who dwell there and have certain advantages over places considered as rural areas where there is lack of such amenities. The current definition of rural and urban areas by the GSS which considers only the population density of a location irrespective of its developmental status should be reviewed.

Currently, if a location has a population of 5000 or more irrespective of availability of amenities like access roads, electricity, good drainage systems, banks etc that location is considered as an urban area. Such categorization does not bring out the actual effects or urban and rural characteristics on responses of surveyed households. Other indicators like availability of electricity, good drainage systems and access roads, organized commercial centers, banks etc should be introduced in the equation to develop an index to determine the urban or rural status of a location and not just the population size of the location. As it is now, research conclusions based on the current definition could be misleading and deceptive.

REFERENCES

Aitkin, M., Anderson, D., & Hinde, J. (1981). Statistical modelling of data on teaching styles. *Journal of the Royal Statistical Society, 144*(4), 419-461. doi: 10.2307/2981826

Anderson, S., Murray, L., & Brownlie, J. (2002). Disciplining Children: Research with parents in Scotland. Edinburgh: Scottish Executive Central Research Unit.

Apt, N. A., Agbenyega, D. L., & Ame, R. K. (2012). Confronting the Challenges: Optimizing Child Rights in Ghana. Children's Rights in Ghana: Reality or Rhetoric? (pp. 3-11). London: Mot Juste Limited.

Ateah, C. A., & Durrant, J. (2005). Maternal use of physical punishment in response to child misbehavior: Implications for child abuse prevention. *Child Abuse & Neglect, 29*(2), 169-185. doi: 10.1016/j.chiabu.2004.10.010

Azen, R., & Walker, C. M. (2011). Categorical data analysis for the behavioural and social sciences. New York, NY: Routledge.

Bagherpour, S., Bahrami, E., Fathi Ashtiani, A., Ahmadi, K., & Ahmadi, A. (2007). The relationship between Parenting styles and mental health and academic success. *Journal of Behavioral Sciences, 1*(1), 33-40.

Baumrind, D. (1991). The Influence of parenting style on adolescent competence and substance use. *The Journal of Early Adolescence, 11*, 56-95.

Bennett, N. (1976). Teaching styles and pupil progress. Harvard University Press. https://books.google.com.gh

Bower-Russa, M., Knutson, J. F., & Winebarger, A. (2001). Disciplinary history, adult disciplinary attitudes, and risk for abusive parenting. *Journal of Community Psychology*, *29*(3), 219-240. doi: 10.1002/jcop.1015

Brody, G., & Flor, D. (1998). Maternal resources, parenting practices, and child competence in rural, single parent African American families. *Child Development*, *69*, 803-816.

Bronfenbrenner, U. (1997). The ecology of human development: Experiments by nature and design. Cambridge, Mass: Harvard University Press.

Bugental, D. B., Johnson, C., New, M., & Sylvester, J. (1998). Measuring parental attributions: Conceptual and methodological issues *Journal of Family Psychology*, *12*, 459-480

Chilman, C. S. (1993). Parental Employment and Child Ca1re Trends: Some Critical Issues and Suggested Policies. https://doi.org/10.1093/sw/38.4.451. *Socical Work*, *38*(4), 451-460.

Coyl-Shepherd, D. D., Roggman, L. A., & Newland, L. A. (2002). Stress, maternal depression, and negative mother-infant interactions in relation to infant attachment. *Infant Mental Health Journal*, *23*(1), 145-163. doi: 10.1002/imhj.10009

Crane, J. (1991). The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing. American Journal of Sociology.

Cummings, E. M., Davies, P. T., & Campbell, S. B. (2000). Developmental psychopathology and family process: Theory, research and clinical implications. New York: Guilford Press.

Dasen, P. R., & Mishra, R. C. (2000). Cross-cultural views on human development in the third millennium. *International Journal of Behavioural Development, 24*, 428-434.

Datta, S., & Satten, G. A. (2008). A Signed-Rank Test for Clustered Data. *Biometrics, 64*(2), 501-. doi: 10.1111/j.1541-0420.2007.00923.x

Deater-Deckard, K., Lansford, J., Dodge, K., Pettit, G., & Bates, J. (2003). 'The development of attitudes about physical punishment: An 8-year longitudinal study'. *Journal of Family Psychology, 17*(3), 351-360.

Delgado, M. (2002). Providing child care for Hispanic families. *Young Children, 35*, 26-32.

Dietz, T. L. (2000). Disciplining children: Characteristics associated with the use of corporal punishment. *Child Abuse & Neglect, 24*, 1529-1542.

Dobson, J. (1977). Dare to Discipline. *Bantam Books isbn = 0-553-22841-2.*

Dodge, K. A. (2002). Mediation, moderation, and mechanisms in how parenting affects children's aggressive behaviour. In J. G. Borkowski, S. L. Ramey, & M. Bristol-Power (Eds.), Parenting and the child's world: Influences on academic, intellectual and social development. Mahwah, NJ: Erlbaum. 215–229.

Domjan, M. (2000). The Essentials of Conditioning and Learning (2nd edition). Belmont, CA: Wadsworth.

Donner, A., & Banting, D. (1988). Analysis of sitespecific data in dental studies. *J Dent Res, 67*, 1392–1395.

Dubowitz, H., Lane, W. G., Semiatin, J. N., & Magder, L. S. (2010). The SEEK model of pediatric primary care: Can child maltreatment be prevented in a low-risk population? Academic Pediatrics, 12(4), 256-268. https://doi.org/10.1016/j.acap.2012.03.005.

Durrant, J. E. (1999). Evaluating the success of Sweden's corporal punishment ban Child Abuse & Neglect. *Child Abuse & Neglect, 23*(5), 435-448. doi: 10.1016/S0145-2134(99)00021-6

Durrett, M. E., O'Bryant, S., & Pennebacker, J. W. (1975). Child-rearing reports of white, black, and Mexican American families. *Developmental Psychology, 11*, 871.

Eamon, M. K., & Zuehl, R. M. (2001). Maternal Depression and Physical Punishment as Mediators of the Effect of Poverty on Socioemotional Problems of Children in Single-Mother Families. *American Journal of Orthopsychiatry, 71*(2), 218-226. doi: 10.1037//0002-9432.71.2.218

Eisenberg, N., & Valiente, C. (2002). 'Parenting and Children's Prosocial and Moral Development'. *In: M.H. Bornstein (ed.), Handbook of Parenting*

5(2), 111-142.

Ellison, C. G., Bartkowski, J. P., & Segal, M. L. (1996). Conservative Protestantism and the parental use of corporal punishment. *Social Forces, 74*, 1003-1028.

English, D. J., Graham, J. C., Newton, R. R., Lewis, L. T., Thompson, R., & Kotch, B. J. e. a. (2008). At-risk and maltreated children exposed to intimate partner aggression/violence: what the conflict lookslike and Its relationship to child outcomes. *Child Maltreatment, 14*(2), 1–15.

Entwisle, B., Mason, W., & Hermalin, A. (1986). The Multilevel Dependence of Contraceptive Use on Socioenomic Development and Family Planning Program Strength. *Demography*, 23(2), 199-216. doi:10.2307/2061616

Feng, S., Kapoor, T. M., Shirai, F., Combs, A. P., & Schreiber, S. L. (1996). Molecular basis for the binding of SH3 ligands with non-peptide elements identified by combinatorial synthesis. Research Support, Non-U.S. Gov't |Research Support, U.S. Gov't, P.H.S. *Chem Biol, 3*(8), 661-670.

Fidell Barbara, G., & Tabachnick Linda, S. (2007). Using multivariate statistics. Boston; Montreal: Pearson/A & B. ISBN 0-205-45938-2. *5*.

Fleming, S. (2012). "How has child discipline changed?". *(http://www.livestrong.com/article/31506-child-discipline-changed/). Demand Media*.

Garcia, C. C., Meyer, E., & Brillon, L. (1995). Ethnic and minority parenting. In M. H. Bornstein (Ed.). *Handbook of parenting: Biology and ecology of parenting. Mahwah, NJ:Erlbaum, 2*, 189-210.

Gershoff, E. T. (2002). Corporal punishment by parents and associated child behaviour and experiences: ameta-analytic and theoretical review. *Psychological Bulletin, 128*, 539–579.

Gershoff, E. T., Grogan-Kaylor, A., Lansford, J. E., Chang, L., Zelli, A., Deater-Deckard, K., & Dodge, K. A. (2015). Parental Discipline Practices in an International Sample: Associations with child behaviours and moderation by perceived normativeness. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2888480).

Ghana News Agency. (2014). Key initiatives in corporate social responsibility: Global Dimension.

Ghana Statistcal Service. (2010). Population and Housing census.

Ghana Statistical Service. (2006). Multiple Indicator Cluster Survey Report.

Ghana Statistical Service. (2011). Multiple Indicator Cluster Survey Report.

Giles-Sims, J., Straus, M. A., & Sugarman, D. B. (1995). 'Child, maternal and family characteristics associated with spanking'. *Family Relations, 44*(2), 170-176.

Goldstein. (1991). Approach to the determinants of urban and rural fertility in Bangladesh. *Asia-Pacific population journal, 12*.

Goldstein, H. (2003a). Multilevel statistical models *London: Edward Arnold*(3).

Goldstein, H. (2003b). Multilevel statistical models. New York. *Distributed in the United States of America by Oxford University Press*(3).

Grogan-Kaylor, A., & Otis, M. D. (2007). The predictors of parental use of corporal punishment. *Family Relations, 56*, 80–91.

Grusec, J. E., & Goodnow, J. J. (1994). 'Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current points of view'. *Developmental Psychology, 30*, 4-19.

Halpenny, A. M., Nixion, E., & Watson, D. (2010). Parents' Perspectives on Parenting Styles and Disciplining Children. Published by Stationery Office, Dublin.

Harrison, A. O., Wilson, M. N., Pine, C. J., Chan, S. Q., & Buriel, R. (1990). Family ecologies of ethnic minority children. *Child Development,, 61*, 347-362.

Heck, R. H., & Thomas, S. L. (2000). An introduction to multilevel modelling techniques. Mahwah, NJ: Erlbaum.

Hedeker, D., & Mermelstein, R. J. (2007). Mixed-effects regression models with heterogeneous variance: Analyzing ecological momentary assessment (EMA) data of smoking.

Hemenway, D., Solnick, S., & Carter, J. (1994). 'Child-rearing Violence'. *Child Abuse and Neglect, 18*(12).

Heo, M., & Leon, A. C. (2005). Comparison of statistical methods for analysis of clustered binary observation. *Statistics in Medicine, 24*(6), 911-923. doi: 10.1002/Sim 1958

Heppner, P., Wampold, B., & Kivlighan, D. (2008). 'Research Design in Counselling'.

Hill, N. E., & Bush, K. (2001). Relations between parenting environment and children's mental health among African American and Euro-American mothers and children. *Journal of Marriage and the Family, 63*, 954-966.

Hoff-Ginsberg, E., & Tardiff, T. (1995). Socioeconomic status and parenting. *In M. H. Bornstein (Ed.), Hand book of parenting: Biology and ecology of parenting Mahwah, NJ: Erlbaum, 2*, 161-188.

Hoffman, L. W. (2002). Methodological issues in studies of SES, parenting, and child development.

Hoffman, M. L. (1983). 'Affective and Cognitive Processes in Moral Internalization'. *In: E.T. Higgins, D. Ruble and W. Hartup (eds.), Social Cognition and Social Development: A Socio-Cultural Perspective. Cambridge: Cambridge University Press*, 236-274.

Holden, G. W. (2002). 'Perspectives on the effects of corporal punishment: Comment on Gershoff'. *Psychological Bulletin, 128*(4), 590-595.

Holden, G. W. (2012). Child Safety, Welfare and Well-being: Issues and Challenges. https://books.google.com.gh

Holden, G. W. (2015a). Discipline and Punishment. http://www.oxfordbibliographies.com/view/document/obo-978019.

Holden, G. W. (2015b). Discipline and Punishment. Retrieved 15/07/15 from http://www.oxfordbibliographies.com/view/document/obo-978019.

Hox, J. J. (2010). Multilevel analysis: Techniques and applications. Mahwah, NJ: Erlbaum.

John, L. (1693). Some Thoughts Concerning Education. http://www.faqs.org/childhood/Co-Fa/Discipline.html.

Jones, K., & Duncan, C. (1995). Specifying and estimating multi-level models of geographical research. *T I Brit Georgroy, 16*, 148-159.

Khouynejad, G., Rajayi, A., & Mohebrad, T. (2006). The relationship between perceived parenting styles and feeling lonely among adolescent girls. *Science and Research in Applied Psychology, 9*(34), 75-92.

Kochanska, G., & Thompson, R. A. (1997). The emergence and development of conscience in toddlerhood and early childhood, in.Grusec J.E and Kuczynski L. (Eds.), Parenting and Children's Internalization of Values: A Handbook of Contemporary Theory. (New York: Wiley, 1997), 53-77.

Kreft, I., & De Leeuw, J. (1998). Introduction to Multilevel Modelling. Sage Publications, London.

Latkin, C. A., Edwards, C.,Davey-Rothwell, M.A.,Tobin, K.E. (2017).The relationship between social desirability bais and self-reports of health, substance use, social network factors among urban substance users in Baltimore, Maryland. Addict Behav. 2017 October: 73: 133-136, doi:10.1016/j.addbeh.2017.05.005.

Lamborn, S. D., Mounts, N. S., Steinberg, L., & Dornbusch, S. M. (1991). Patterns of competence and adjustment among adolescents from authoritative, authoritarian, indulgent, and neglectful families. *Child Development*, *62*(1049-1065).

Lareau, A. (2002). "Invisible Inequality: Social Class and Childrearing in Black Families and White Families". *American Sociological Review*, *67*(5), 747–776. doi: 10.2307/3088916. JSTOR 3088916

LaReau, A. (2003). Unequal childhoods: Class, race, and family life. Berkeley, CA. *University of California Press*.

Larocque, D., Haataja, R., Nevalainen, J., & Oja, H. (2010). Two sample tests for the nonparametric Behrens–Fisher problem with clustered data. *Journal of Nonparametric S tatistics*, *22*(6), 755-771. doi: 10.1080/10485250903469728

Lau, A. S. (2010). Physical discipline in Chinese American immigrant families: An adaptive culture perspective. *Cultural Diversity and Ethnic Minority Psychology*. *http://dx.doi.org/10. 1037/a0018667*, *16*(3), 313-322.

Lerner, R. M., Rothbaum, F., Boulos, S., & Castellino, D. R. (2002). 'Developmental Systems Perspective on Parenting'. *In: M.H. Bornstein (ed.), Handbook of Parenting. Biology and Ecology of Parenting. Mahwah, NJ: Lawrence Erlbaum Associates*, *2*(2), 315-344.

Litiere, S., Alonso, A., & Molenberghs, G. (2007). Type I and type II error under random-effects misspecification in generalized linear mixed models. *Biometrics, 63*(4), 1038-1044.

Luke, D. A. (2004). Multilevel modelling (3. repr. ed.). Thousand Oaks, CA: Sage. ISBN 978-0-7619-2879-9.

Maccoby, E. E. (1992). The role of parents in the socialization of children: An historical overview. *Developmental Psychology, 28*, 1006–1017.

McLoyd, V. C., & Jayaratne, T. E. (1994). Unemployment and Work Interruption among African American Single Mothers: Effects on Parenting and Adolescent Socioemotional Functioning. *Child Development, 65*(2).

McKeown, D. (2006). *Early child development family abuse prevention project actions to prevent physical punishment of children: Staff report Toronto: City of Tronto* Retrievedfrom http://www.toronto.ca/legdocs/2006/agendas/committees/h1/h1060410/it003.

pdf.

Moerbeek, M. (2004). The Consequence of Ignoring a Level of Nesting in Multilevel Analysis *Multivariate behavioral research* *39*(1), 1-.

Murphy-Cowan, T., & Stringer, M. (1999). 'Physical Punishment and the Parenting Cycle: A Survey of Northern Irish Parents'. *Journal of Community and Applied Social Psychology, 9*, 61-71.

MyJoyOnline. (2010). "Scarred with Whips: The agony of Osu Children's Home inmates".

Nancy, E. H. (2006). Disentangling ethnicity, socioeconomic status and parenting: Interactions, influences and meaning. *Vulnerable Children and Youth Studies, 1*(1), 114-124. doi: 10.1080/17450120600659069.

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 *1991, 59*(1), 25-35.

Neuman. (2000). Sociology and psychology in public health. Oxford Textbook of global public health. https://books.google.com.gh.

Nobes, G., Smith, M., Upton, P., & Heverin, A. (1999). 'Physical punishment by mothers and fathers in British homes'. *Journal of Interpersonal Violence, 14*(8), 887-902.

O'Connell, A. A., Goldstein, J., Rogers, H., & Peng, C. Y. (2008). Multilevel logistic models for dichotomous and ordinal data. https://www.researchgate.net/publication/281423398_multilevel_logistics.

O'Connell, A. A., Goldstein, J., Rogers, H. J., & Peng, C. (2008). Multilevel logistic models for dichotomous and ordinal data. https://www.researchgate.net/publication / 28142339.

Oman, S. D., & Zucker, D. M. (2001). Modelling and generating correlated binary variables. *Biometrika, 88*(1), 287-290.

Orme, N. (2001). Medieval Children. Yale University Press, 2001. History – 387 pages.

Cole, P.M., Michel, M.K., & Teti, L.O. (1994). The development of emotion regulation and dysregulation: A clinical perspective. *Monographs of the Society for Research in Child Development, 59*(2/3, Serial No. 240), 73–102.

Patterson, G. R. (1982). A social learning approach. *Coercive family process. Eugene, OR: Castalia, 3*.

Pinderhughes, E., Dodge, K. A., & Bates, J. E. (2000). Discipline Responses: Influences of Parents' Socioeconomic Status, Ethnicity, Beliefs About Parenting, Stress, and Cognitive–Emotional Processes. *Journal of Family Psychology 14*(3), 380-400.

Pollock Linda, A. (1983). Forgotten children. *Parent-child relations from 1500 to 1900. Cambridge University Press. ISBN 0-521-25009-9*.

Rahmani, F., Sayyed, F. N., Rezayi, M., Sedaghat, K., & Fathi, A. E. (2005). Relationship between parenting style and behavior problems of students. *Iran Psychiatry and Clinical Psychology, 12*(4), 365-370.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models. *Applications and data analysis methods. Thousand Oaks; Sage*.

Reich, S. M., Penner, E. K., Duncan, G. J., & Auger, A. (2012). Using baby books to change new mothers' attitudes about corporal punishment. Child abuse Neglect, 36(2), 108-117. https//doi.10.1016/j.chiabu.2011.09.017.

Rogoff, B. (2003). The Cultural Nature of Human Development. Oxford University Press.

Rosner, B., Glynn, R. J., & Lee, M. (2006). Extension of the rank sum test for clustered data: two-group comparisons with group membership defined at the subunit level. *Biometrics*, *62*, 1251–1259.

Rosner, B., Glynn, R. J., & Lee, M. L. T. (2003). Incorporation of clustering effects for the Wilcoxon rank sum test. *A large-sample approach. Biometrics, 59,* 1089 – 1098.

Rosner, B., & Grove, D. (1999). Use of the Mann–Whitney U-test for clustered data. *Stat Med*, *18*, 1387–1400.

Rountree, P. W., & Land, K. C. (1996). Burglary Victimization, Perceptions of Crime Risk, and Routine Activities: A Multi- level Analysis across Seattle Neighborhoods and Census Tracts. *Journal of Research in Crime and Delinquency, 33*, 147-180.

Somnath, D., & Satten, G.A (2012). Rank-sum tests for clustered data. *J Am Stat Assoc 100*, 908 –915.

Sadoughi, Z., Vafayi, M., & Rasoulzadeh. (2007). Relationship between sensory processing sensitivity with parenting practices and indicators of psychological pathology. *Advances in Cognitive Science, 9*(1), 23-32.

Sadrosadat, J., Shams, H., & Emamipour, S. (2005). Comparison of parenting practices and family functioning in irresponsible families with behavioral disorder and normal. *Shahrekord University of Medical Sciences, 7*(2), 43-48.

Salkind, S. B., & Green, N. J. (2004). Using SPSS for Windows and Macintosh : analyzing and understanding data. . *Upper Saddle River, NJ: Pearson Education. ISBN 0-13-146597-X.*(4).

(SAS Reference materials/SAS 9.4/ Logistic procedure. https://www.uidaho.edu., n.d.)

Sayyed, M. P., NaadAli, H., & Ghanbari, S. (2008). The relationship between parenting styles and externalizing symptoms in 7 to 9 years children. *Journal of Family Research*, *4*(13), 37-49.

Smith, A. B., Gollop, M., Taylor, N. J., & Marshall, K. (2005). The Discipline and Guidance of Children: A Summary of Research. Dunedin and Wellington, NZ: Children's Issues Centre and Office of the Children's Commissioner.

Steinberg, L., Lamborn, S. D., Dornbusch, S. M., & Darling, N. (1992). Impact of parenting practices on adolescent achievement: authoritative parenting, school involvement, and encouragement to succeed. *Child Development*, *63*, 1266-1281.

Stirateli, R., Laird, N., & Ware, J. H. (1984). Randon-effects models for serial observation with binary response. *Biometrics, 40*(4), 961-971.

Stoolmiller, M., Snyder, J., Wilson, M. N., & Yamamoto, M. (2003). Child Anger Regulation, Parental Responses to Children's Anger Displays, and Early Child

Antisocial Behavior. Social Development 12(3):335 – 360 DOI: 10.1111/1467-9507.00237 *Social Development, 12*(3), 335-360. doi: 10.1111/1467-9507.00237

Straus, M. A. (1991). 'Discipline and Deviance: Physical punishment of children and violence and other crime in adulthood'. *Social Problems, 38*, 133-154.

Straus, M. A. (2001). Beating the devil out of them: corporal punishment in American families and itseffects on children. *New Brunswick: Transaction Publishers*(2).

Straus, M. A., Douglas, E. M., & Mederios, R. A. (2014). The Primordial Violence: Spanking Children, Psychological Development, Violence, and Crime. New York: Routledge.

Straus, M. A., & Stewart, J. H. (1999). 'Corporal Punishment by American Parents: National data on prevalence, chronicity, severity and duration in relation to child and family characteristics'. *Clinical Child and Family Psychology Review, 2*(2), 55-70.

Tranmer, M., & Steel, D. G. (2001). Ignoring a level in a multilevel model: Evidence from UK census data. https://doi.org/10.1068/a3317. *Environment and Planning A, 33*(5).

Turner, H. A., & Muller, P. A. (2004). Long-term effects of child corporal punishment on depressive symptons in young adults: Potential moderators and mediators. *Jounrnal of Family Issues, 25*, 761 – 782.

Twum-Danso, A. (2010). Children's Perceptions of Physical Punishment in Ghana, Nuffield Foundation.

Twum-Danso, A. (2012). 'From Central to Marginal? Changing Perceptions of Kinship Fosterage in Ghana'. *In Journal of Family History, Sage Publications*.

UNICEF, (2014). Ending Violence against children: Six Strategies for Action. United Nations Children's Fund.

Van, G. W., Legerstee, J. S., & Ferdinand, R. F. (2009). The role of perceived parenting in familial aggregation of anxiety disorders in children. *Journal of Anxiety Disorders, 23*(1), 46-53. doi: 10.1016/j.janxdis.2008.03.014.

Vega, W. A. (1980). Hispanic families in the 1980s: A decade of research. *Journal of Marriage and the Family*, *52*, 1015-1024.

Verhoeven, M., Bögels, S. M., & Van der Bruggen, C. C. (2011). Unique Roles of Mothering and Fathering in Child Anxiety; Moderation by Child's Age and Gender. *Journal of Child Family Studies, 21*(2), 331-343.

Virmani, R., Burke, A. P., Farb, A., & Kolodgie, F. D. (2006). The Pathology of Vulnerable Plaque. *Journal of the American College of Cardiology*, *47*(8). doi: 10.1016/j.jacc.2005.10.065

Willeumier, K., Pulst, S., & Scheizer, F. (2006). Proteasome Inhibition Triggers Activity-Dependent Increase in the Size of the Recycling Vesicle Pool in Cultured Hippocampal Neurons. *Journal of the Society for Neuroscience, 26*(44), 11333-11341.

Wilson, P. M., Rogers, W. T., Rodgers, & Wild, T. C. (2002). The Psychological Need Satisfaction in Exercise Scale *Journal of Sport & Exercise Psychology, 28*(231-251).

Wissow, L. S. (2001). Ethnicity, income, and parenting contexts of physical punishment in a national sample of families with young children. *Child Maltreatment*, *6*, 118-129.

Wissow, L. S. (2002). Child discipline in the first three years of life in Halfon N., K.T. McLearn and M.A. Schuster (eds.) Child Rearing in America: Challenges Facing Parents with Young Children. New York: Cambridge University Press (pp.146–177).

Xu, Y., Farver, J. M., Bhadaha, B. R., Narang, S., & Lieber, E. (2007). Ethnic Identity, Acculturation, Parenting Beliefs, and Adolescent Adjustment: A Comparison of Asian Indian and European American Families. *Merrill-Palmer Quarterly, 53*(2), 184-215. doi: 10.1353/mpq.2007.0010

Yin, R. K. (1994). Case Study Research – Design and Methods, Applied Social Research, 5th edn, Methods Series, Sage Publications, Newbury Park.

Zikmund, W. G. (2000). Business Research Methods. 6th edn, The Dry den Press, Chicago.

APPENDICES

APPENDIX A

BINARY RESULTS

Table A1: Background information on respondents

Background characteristics of respondents	N	Percentage (%)	Child needs physical discipline (%)	
			Yes	**No**
Area				
Urban households	2,683	32.9	49.0	51.0
Rural households	5,473	67.1	50.9	49.1
Region				
Western	473	5.8	36.2	63.8
Central	1,192	14.6	56.7	43.3
Greater Accra	511	6.3	42.1	57.9
Volta	491	6.0	60.3	39.7
Eastern	457	5.6	68.1	31.9
Asante	568	7.0	54.9	45.1
Brong Ahafo	480	5.9	33.1	66.9
Northern	1,615	19.8	60.9	39.1
Upper East	1,158	14.2	36.8	63.2
Upper West	1,211	14.8	45.7	54.3
Education of household head				
None	4,259	52.2	51.5	48.5
Primary	1,104	13.5	49.6	50.4
Middle/JSS	2,196	26.9	50.6	49.4
Secondary +	597	7.3	42.2	57.8
Ethnicity of household head				
Akan	2,404	29.5	50.5	49.5
Ga/Dangme	343	4.2	57.7	42.3
Ewe	745	9.1	54.2	45.8
Guan	317	3.9	38.5	61.5
Gruma	540	6.6	56.5	43.5
Mole Dagbani	2,981	36.5	50.5	49.5
Grusi	479	5.9	43.2	56.8
Mande	148	1.8	38.5	61.5
Non-Ghanaian	124	1.5	48.4	51.6
Others	75	0.9	38.7	61.3
Wealth index quintile				
Poorest	3,336	40.9	50.1	49.9
Second	1,656	20.3	53.3	46.7

CONTINUATION OF TABLE A1

Middle	1,249	15.3	54.4	45.6
Fourth	1,066	13.1	49.7	50.3
Richest	849	10.4	39.8	60.2
Sex of household head				
Male	6,190	75.9	50.4	49.6
Female	1,966	24.1	50.0	50.0
Religion of household head				
Christian	4,389	53.8	48.3	51.7
Moslem	2,104	25.8	54.3	45.7
Traditional	1,147	14.1	49.3	50.7
other	516	6.3	53.3	46.7
Age group of head				
15-21 years	40	0.5	27.5	72.5
22-40 years	2,832	34.7	50.7	49.3
41-60 years	3,615	44.3	49.8	50.2
61+years	1,669	20.5	51.2	48.8
Number of children (2-14) in household				
1 child	2,357	28.9	45.6	54.4
2 children	2,160	26.5	51.7	48.3
3+	3,639	44.6	52.5	47.5
Child needs physical discipline				
Yes	4,103	50.3		
No	4,053	49.7		

Table A2: Frequency distribution of household's approval of physical discipline by predictor.

Variables (chi-square and p-value)	level	Approval of physical discipline		
		Yes	No	Total
Education of household head Chi-square =18.22 p<0.0004	None	2192	2067	4259
		26.88%	25.34%	52.22%
	Primary	548	556	1104
		6.72%	6.82%	13.54%
	Middle/JSS	1111	1085	2196
		13.62%	13.30%	26.92%
	Secondary +	252	345	597
		3.09%	4.23%	7.32%
Ethnicity of household head Chi-squared = 60.32 p<0.0001	Akan	1215	1189	2404
		14.90%	14.58%	29.48%
	Ga/Dangme	198	145	343
		2.43%	1.78%	4.21%
	Ewe	404	341	745
		4.95%	4.18%	9.13%
	Guan	122	195	317
		1.50%	2.39%	3.89%
	Gruma	305	235	540
		3.74%	2.88%	6.62%
	Mole Dagbani	1506	1475	2981
		18.46%	18.08%	36.55%
	Grusi	207	272	479
		2.54%	3.33%	5.87%
	Mande	57	91	148
		0.70%	1.12%	1.81%
	Non-Ghanaian	60	64	124
		0.74%	0.78%	1.52%
	Others	29	46	75
		0.36%	0.56%	0.92%

CONTINUATION OF TABLE A2:

Wealth Index Chi-square = 52.17 p<0.0001	Poorest	1672	1664	3336
		20.50%	20.40%	40.90%
	Second	883	773	1656
		10.83%	9.48%	20.30%
	Middle	680	569	1249
		8.34%	6.98%	15.31%
	Fourth	530	536	1066
		6.50%	6.57%	13.07%
	Richest	338	511	849
		4.14%	6.27%	10.41%
Gender of household head Chi-square = 0.10 p<0.7550	Male	3120	3070	6190
		38.25%	37.64%	75.90%
	Female	983	983	1966
		12.05%	12.05%	24.10%
Religion of household head Chi-square = 23.07 p<0.0001	Christian	2119	2270	4389
		25.98%	27.83%	53.81%
	Moslem	1143	961	2104
		14.01%	11.78%	25.80%
	Traditional	566	581	1147
		6.94%	7.12%	14.06%
	other	275	241	516
		3.37%	2.95%	6.33%
Age-group of household head chi-square = 9.41 p<0.0243	15-21 years	11	29	40
		0.13%	0.36%	0.49%
	22-40 years	1436	1396	2832
		17.61%	17.12%	34.72%
	41-60 years	1801	1814	3615
		22.08%	22.24%	44.32%
	61years	855	814	1669
		10.48%	9.98%	20.46%
Number of children in household chi-square = 30.17 p<0.0001	1 child	1074	1283	2357
		13.17%	15.73%	28.90%
	2 children	1117	1043	2160
		13.70%	12.79%	26.48%
	3+	1912	1727	3639
		23.44%	21.17%	44.62%

MODELLING TABLES FOR BINARY LOGIT MODEL

Table A3: Level-1 fixed effects estimates and standard errors for binary logit model

Effect	Category	Estimate	Std Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept		-0.4121	0.339	37	-1.22	0.2317	0.05	-1.099	0.2747
Education of Head	Middle/JSS	-0.0072	0.2559	8093	-0.03	0.9775	0.05	-0.5088	0.4944
	None	0.0808	0.257	8093	0.31	0.7533	0.05	-0.423	0.5846
	Primary	-0.0866	0.2627	8093	-0.33	0.7416	0.05	-0.6015	0.4282
	Secondary +	0	.	.	.	.	.	.	.
Ethnicity of Head	Akan	0.6635	0.2635	8093	2.52	0.0118	0.05	0.147	1.1801
	Ewe	0.492	0.2721	8093	1.81	0.0706	0.05	-0.0413	1.0254
	Ga/Dangme	0.8419	0.285	8093	2.95	0.0031	0.05	0.2831	1.4006
	Gruma	0.5271	0.2667	8093	1.98	0.0482	0.05	0.0043	1.05
	Grusi	0.4887	0.264	8093	1.85	0.0642	0.05	-0.0288	1.0062
	Guan	-0.2377	0.2759	8093	-0.86	0.3889	0.05	-0.7785	0.3031
	Mande	0.488	0.3039	8093	1.61	0.1083	0.05	-0.1077	1.0837
	Mole Dagbani	0.685	0.2491	8093	2.75	0.006	0.05	0.1967	1.1734
	Non-Ghanaian	0.2545	0.3091	8093	0.82	0.4103	0.05	-0.3513	0.8603
	Others	0	.	.	.	.	.	.	.
Sex of Head	Female	0.0521	0.05799	8093	0.9	0.3689	0.05	-0.06157	0.1658
	Male	0	.	.	.	.	.	.	.
Religion of Head	Christian	-0.1597	0.09857	8093	-1.62	0.1052	0.05	-0.353	0.03351
	Moslem	0.0725	0.1102	8093	0.66	0.5105	0.05	-0.1435	0.2886
	Traditional	-0.0841	0.1123	8093	-0.75	0.4539	0.05	-0.3042	0.136
	other	0	.	.	.	.	.	.	.
AgeGroup of Head	15-21 years	**-1.0438**	0.3702	8093	-2.82	**0.0048**	0.05	-1.7696	-0.3181
	22-40 years	0.0593	0.0669	8093	0.89	0.3754	0.05	-0.0718	0.1904
	41-60 years	-0.0084	0.06325	8093	-0.13	0.8944	0.05	-0.1324	0.1156
	61years	0	.	.	.	.	.	.	.
Children (2-14) in household	1 child	-0.2305	0.05804	8093	-3.97	<.0001	0.05	-0.3442	-0.1167
	2 children	0.0311	0.0576	8093	0.54	0.5895	0.05	-0.0819	0.144
	3+	0	.	.	.	.	.	.	.
Wealth index of Household	Fourth	-0.2124	0.08624	8093	-2.46	0.0138	0.05	-0.3814	-0.04331
	Middle	-0.0466	0.0796	8093	-0.59	0.5581	0.05	-0.2027	0.1095
	Poorest	0.0557	0.0715	8093	0.78	0.4361	0.05	-0.0844	0.1957
	Richest	-0.5165	0.1046	8093	-4.94	<.0001	0.05	-0.7216	-0.3114
	Second	0	.	.	.				

Table A4: Odds Ratio Estimates for multilevel binary logit model

Comparison	Estimate	DF	95% Confidence Limits	
Head_Education None vs Middle/JSS	1.092	8093	0.68	1.753
Head_Education Primary vs Middle/JSS	0.924	8093	0.57	1.498
Head_Education Secondary + vs Middle/JSS	1.007	8093	0.61	1.663
Head_Ethnicity Ewe vs Akan	0.842	8093	0.683	1.039
Head_Ethnicity Ga/Dangme vs Akan	1.195	8093	0.925	1.545
Head_Ethnicity Gruma vs Akan	0.872	8093	0.678	1.122
Head_Ethnicity Grusi vs Akan	0.84	8093	0.643	1.097
Head_Ethnicity Guan vs Akan	0.406	8093	0.309	0.534
Head_Ethnicity Mande vs Akan	0.839	8093	0.557	1.263
Head_Ethnicity Mole Dagbani vs Akan	1.022	8093	0.833	1.254
Head_Ethnicity Non-Ghanaian vs Akan	0.664	8093	0.448	0.985
Head_Ethnicity Others vs Akan	0.515	8093	0.307	0.863
Head_Sex Male vs Female	0.949	8093	0.847	1.063
Head_Religion Moslem vs Christian	1.261	8093	1.096	1.452
Head_Religion Traditional vs Christian	1.079	8093	0.927	1.254
Head_Religion other vs Christian	1.173	8093	0.967	1.423
Head_AgeGroup 22-40 years vs 15-21 years	3.014	8093	1.464	6.201
Head_AgeGroup 41-60 years vs 15-21 years	2.816	8093	1.369	5.793
Head_AgeGroup 61years vs 15-21 years	2.84	8093	1.374	5.868
Num_Children 2 children vs 1 child	1.299	8093	1.149	1.468
Num_Children 3+ vs 1 child	1.259	8093	1.124	1.411
HHold_Windex Middle vs Fourth	1.18	8093	0.993	1.402
HHold_Windex Poorest vs Fourth	1.307	8093	1.092	1.565
HHold_Windex Richest vs Fourth	0.738	8093	0.604	0.902
HHold_Windex Second vs Fourth	1.237	8093	1.044	1.464

Table A5: Distance between 95% confidence interval limits for estimated parameter values with increasing simulated sample size

	N=1000	N=2000	N=3000	N=4000	N=5000
Intercept	1.0369	0.7189	0.5918	0.5073	0.4528
Age group of head	1.0004	0.698	0.5713	0.4869	0.4417
Ethnicity	1.0057	0.69788	0.573	0.4923	0.4429
No. of Children	0.9957	0.6969	0.5699	0.4871	0.4375
Wealth Index	0.985	0.6895	0.5729	0.4885	0.4368

Table A6: Standard error of estimated parameter values with increasing simulated sample sizes.

True parameter values (SE)	1000	2000	3000	4000	5000
Intercept	0.2645	0.1834	0.151	0.1294	0.1155
Age group of head	0.2552	0.1781	0.1457	0.1242	0.1127
Ethnicity	0.2566	0.178	0.1462	0.1256	0.113
No. of Children	0.254	0.1778	0.1454	0.1243	0.1116
Wealth Index	0.2513	0.1759	0.1462	0.1246	0.1114

APPENDIX B

MULTINOMIAL RESULTS

Table B1: Model information for the multilevel multinomial outcome

Response Variable	Type of Discipline
Response Distribution	Multinomial
Link Function	Generalized Logit
Variance Matrix Blocked By	Region
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment
Number of Observations Read	8156
Number of Observations Used	8156

Response profile

Discipline method	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Non-physical methods	1877	23.01	1877	23.01
Psychological methods	2153	26.40	4030	49.41
Physical methods	625	7.66	4655	57.07
Non-physical and Psychological methods	2602	31.90	7257	88.98
All methods	899	11.02	8156	100.00

In modelling category probabilities, Type of Discipline = Random use of “all three methods” served as the reference category.

BIVARIATE TABLES FOR USE OF DISCIPLINE METHODS

Table B2: Type of discipline method used by region

Region	Type of discipline method					Total (%)
	Non-physical methods %	Psychological methods %	Physical methods %	Non-physical and Psychological methods %	All methods %	
Western	1.63	1.21	0.29	1.94	0.72	473 (5.80%)
Central	4.19	2.99	0.65	3.98	2.80	1,192 (14.62%)
Greater Accra	1.31	1.28	0.45	1.88	1.35	511 (6.27%)
Volta	1.15	1.61	0.59	1.67	1.01	491 (6.02%)
Eastern	1.20	1.25	0.50	1.79	0.86	457 (5.6%)
Asante	1.30	1.85	0.33	2.43	1.05	568 (6.96%)
Brong Ahafo	1.25	1.18	0.38	2.48	0.60	480 (5.89%)
Northern	4.33	6.20	2.10	6.33	0.85	1,615 (19.8%)
Upper East	2.21	5.21	1.25	4.67	0.86	1,158 (14.2%)
Upper West	4.44	3.62	1.12	4.74	0.93	1,211 (14.85%)
Total	1877	2153	625	2602	899	8156
	23.01%	26.4%	7.66%	31.9%	11.02%	100%

Table B3: Frequency distribution of household characteristics by choice of discipline type

Variables (chi-square and p-value)		Non-physical methods %	Psychological methods %	Physical methods %	Non-physical and psychological method %	All methods (%)
Wealth Index Chi-square=230.63 p<0.0001	Poorest	9.31	12.16	3.73	13.27	2.44
	Second	4.77	5.10	1.66	6.44	2.34
	Middle	3.92	3.51	0.91	4.90	2.07
	Fourth	2.91	3.05	0.88	4.16	2.07
	Richest	2.11	2.57	0.49	3.14	2.1
Education of household head Chi-square=1164.92 p<0.0001	None	12.21	16.87	5.37	17.26	0.5
	Primary	2.76	3.08	0.8	4.14	2.76
	Middle/JSS	7.04	3.79	1.21	8.73	6.15
	Secondary +	1.01	2.66	0.28	1.77	1.61
Ethnicity of household head Chi-squared= p<0.0001	Akan	7.36	6.06	1.43	9.51	5.11
	Ga/Dangme	1.04	0.99	0.4	1.04	0.72
	Ewe	1.97	2.32	0.64	2.67	1.53
	Guan	0.86	0.86	0.44	1.32	0.4
	Gruma	1.51	1.95	0.66	2.04	0.47
	Mole Dagbani	7.97	11.39	3.16	12.03	2.00
	Grusi	1.43	1.64	0.59	1.78	0.43
	Mande	0.25	0.61	0.13	0.71	0.11
	Non-Ghanaian	0.37	0.39	0.12	0.48	0.16
	Others	0.26	0.18	0.07	0.32	0.09
Gender of household head Chi-square=4.64 p<0.3258	Male	17.58	20.35	5.85	23.79	8.33
	Female	5.43	6.04	1.81	8.12	2.7
Religion of household head Chi-square=212.067 p<0.0001	Christian	12.84	12.78	3.32	16.82	8.06
	Moslem	5.90	7.69	2.50	8.34	1.37
	Traditional	2.92	4.27	1.36	4.65	0.87
	other	1.36	1.67	0.48	2.10	0.72
Age-group of household head chi-square=75.2124 p<0.0001	15-21 years	0.09	0.10	0.04	0.16	0.11
	22-40 years	7.58	8.86	2.66	10.96	4.66
	41-60 years	10.05	11.82	3.11	14.31	5.03
	61years	5.30	5.62	1.85	6.47	1.23
Number of children in household chi-square=25.7845 p<0.0011	1 child	6.68	7.16	1.90	9.37	3.79
	2 children	5.87	7.01	2.07	8.75	2.77
	3+	10.46	12.22	3.69	13.78	4.46

MODELLING TABLES FOR MULTINOMIAL LOGIT MODEL

Table B4: Model M4 - Effects of significant covariates on type of discipline method

Significant Covariates	Non Physical and Psychological methods		Non Physical methods		Physical methods		Psychological methods	
	estimate	P value	estimate	P value	estimate	P value	estimate	P value
Number of Children								
1 Child	0.0825	0.4075	0.02124	0.8379	-0.1194	0.3869	-0.0179	0.8643
2 Children	0.2540	**0.0147**	0.1395	0.1990	0.1841	0.1799	0.1711	0.1149
3 + Children	0	.	0	.	0	.	0	.
Religion of Head								
Christian	0.0478	0.7798	0.2200	0.2236	0.0817	0.7273	0.0996	0.5804
Moslem	0.3980	0.0583	0.6047	**0.0059**	0.6687	**0.0132**	0.4937	**0.0238**
Traditional	0.0666	0.7579	0.0405	0.8583	0.2138	0.4384	0.0630	0.7786
Other	0	.	0	.	0	.	0	.
Education level of Head								
Middle/JSS	0.2394	0.1042	0.5585	**0.0008**	0.06131	0.8193	-1.05	**<0.0001**
None	3.4043	**<0.0001**	3.6031	**<0.0001**	3.9732	**<0.0001**	2.9423	**<0.0001**
Primary	0.2974	0.0809	0.4528	0.0176	0.3886	0.1843	-0.4711	0.0054
Secondary +	0	.	0	.	0	.	0	.

Table B4A: Odd ratio estimates of significant covariates

Significant Covariates	Non Physical and Psychological methods		Non Physical methods		Physical methods		Psychological methods	
	estimate	95% CI	estimate	95% CI	estimate	95% CI	estimate	95% CI
Number of Children								
1 Child	1.09	(0.893, 1.320)	1.02	(0.833, 1.252)	0.887	(0.677, 1.163)	0.982	(0.800, 1.206)
2 Children	1.29	(1.051, 1.581)	1.15	(0.929, 1.422)	1.202	(0.919, 1.573)	1.187	(0.959, 1.468)
3 + Children								
Religion of Head								
Christian	1.05	(0.750, 1.467)	1.25	(0.874, 1.776)	1.085	(0.686, 1.717)	1.105	(0.776, 1.573)
Moslem	1.49	(0.986, 2.248)	1.83	(1.190, 2.815)	1.952	(1.150, 3.312)	1.638	(1.068, 2.514)
Traditional	1.07	(0.700, 1.632)	1.04	(0.668, 1.624)	1.238	(0.721, 2.127)	1.065	(0.686, 1.652)
Other								
Education level of Head								
Middle/JSS	1.27	(0.952, 1.696)	1.75	(1.261, 2.423)	1.063	(0.628, 1.799)	0.350	(0.262, 0.468)
None	30.09	(19.492, 46.459)	36.71	(23.148, 58.221)	53.154	(29.010, 97.389)	18.960	(12.353, 29.102)
Primary	1.35	(0.964, 1.880)	1.57	(1.082, 2.285)	1.475	(0.831, 2.618)	0.624	(0.448, 0.870)
Secondary +								

Table B5: Covariance parameter estimates for region of household

	Non-physical and psychological aggression methods		Non-physical methods		Physical methods		Psychological aggression methods	
Region	Estimate	P-value	Estimate	P-value	Estimate	P-value	Estimate	P-value
Ashanti	0.0875	0.1100	-0.1094	0.1172	-0.0933	0.1650	0.2276	0.1276
Brong Ahafo	0.2829	0.0213	0.0204	0.8636	0.0088	0.9558	-0.0428	0.7473
Central	-0.2969	**0.0067**	0.0633	0.5463	-0.3207	0.0667	-0.1910	0.1140
Eastern	0.0552	0.6277	-.0.0545	0.6471	0.2457	0.1711	0.0440	0.7374
+6Accra	-0.0900	0.4466	-0.1112	0.3637	0.0527	0.7488	-0.2848	**0.0469**
Northern	0.0458	0.6739	0.0249	0.8229	0.0950	0.4935	0.1094	0.3656
Upper East	0.0472	0.6864	-0.1447	0.2311	0.0781	0.6001	0.3038	**0.0226**
Upper West	-0.0891	0.6457	0.2698	**0.0484**	-0.2327	0.1558	-0.2590	0.0549
Volta	-0.1363	0.2788	-0.1750	0.2020	0.2348	0.2135	0.0908	0.5110
Western	0.0864	0.4545	0.2154	0.0939	-0.0013	0.8260	-0.0013	0.9922

Table B5A: Bootstrap estimates of the significant parameters sampled from 20 repetitions using 50% (4,078 observations) of overall data

	Education level of household head	Religion of household head	Number of children in household
True parameter estimate	-0.2261	-0.00948	-0.0377
Bootstrap estimated coefficients of significant covariates			
Sample No. 1	-0.2173	-0.0055	-0.0415
Sample No. 2	-0.2653	0.0263	-0.0696
Sample No. 3	-0.2537	-0.0329	-0.017
Sample No. 4	-0.2467	0.0032	-0.0570
Sample No. 5	-0.1951	0.0656	-0.0914
Sample No. 6	-0.2688	-0.0268	-0.0215
Sample No. 7	-0.1805	0.0074	-0.0532
Sample No. 8	-0.2190	0.0051	-0.0591
Sample No. 9	-0.2631	-0.0466	-0.0060
Sample No. 10	-0.1955	-0.0002	-0.0639
Sample No. 11	-0.2309	0.0007	-0.0689
Sample No. 12	-0.2171	-0.0033	-0.0612
Sample No. 13	-0.2256	-0.0125	-0.0825
Sample No. 14	-0.2580	0.0009	-0.0407
Sample No. 15	-0.2195	0.0287	0.0077
Sample No. 16	-0.2545	0.0059	-0.0046
Sample No. 17	-0.1992	-0.0325	-0.0302
Sample No. 18	-0.1587	0.0054	-0.0297
Sample No. 19	-0.2046	-0.0173	-0.0438
Sample No. 20	-0.2549	0.0222	-0.0960
$\hat{x}$	-0.2264	-0.0003	-0.0465

Table B6: Classification table for multinomial logit model without predictors

Observed	Predicted					
	Non physical methods	Psychological methods	Physical methods	Non physical and Psychological methods	All methods	Percent Correct
Nonphysical methods	337	412	0	974	154	18.0%
Psychological methods	214	776	0	987	176	36.0%
Physical methods	66	235	0	295	29	.0%
Non physical and Psychological methods	286	686	0	1473	157	56.6%
All methods	127	68	0	526	178	19.8%
Overall Percentage	12.6%	26.7%	.0%	52.2%	8.5%	33.9%

APPENDIX C
EXPLORATORY DATA ANALYSIS RESULTS

Distribution and probability plots for variables

Table C1: Variable statistics - Physical punishment is needed

Moments			
N	8156	Sum Weights	8156
Mean	1.496	Sum Observations	12209
Std Deviation	0.500	Variance	0.2500
Skewness	0.012	Kurtosis	-2.000
Mean	1.496	Std Deviation	0.500
Median	1.000	Variance	0.250

Frequency			
Response	Count	Cell (%)	Cum (%)
Yes, approve physical discipline	4103	50.3	50.3
No, do not approve	4053	49.7	100.0

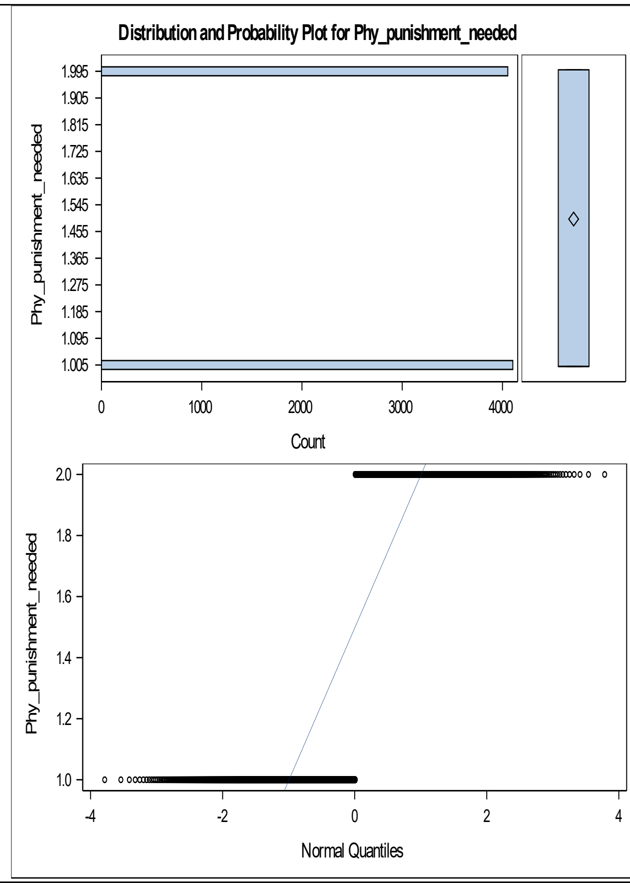


Table C2: Variable statistics - Type of discipline method

Moments			
N	8156	Sum Weights	8156
Mean	2.815	Sum Observations	22961
Std Deviation	1.3822	Variance	1.9105
Skewness	0.0609	Kurtosis	-1.423
Mean	2.815	Std Deviation	1.382
Median	3.000	Variance	1.910
Frequency			
Value	Count	Cell (%)	
1.Nonphysical methods	1877	23.0	23.0
2.Psychological	2153	26.4	49.4
3.Physical methods	625	7.7	57.1
4.Non-phy & Psy methods	2602	31.9	89.0
5.All methods	899	11.0	100.0

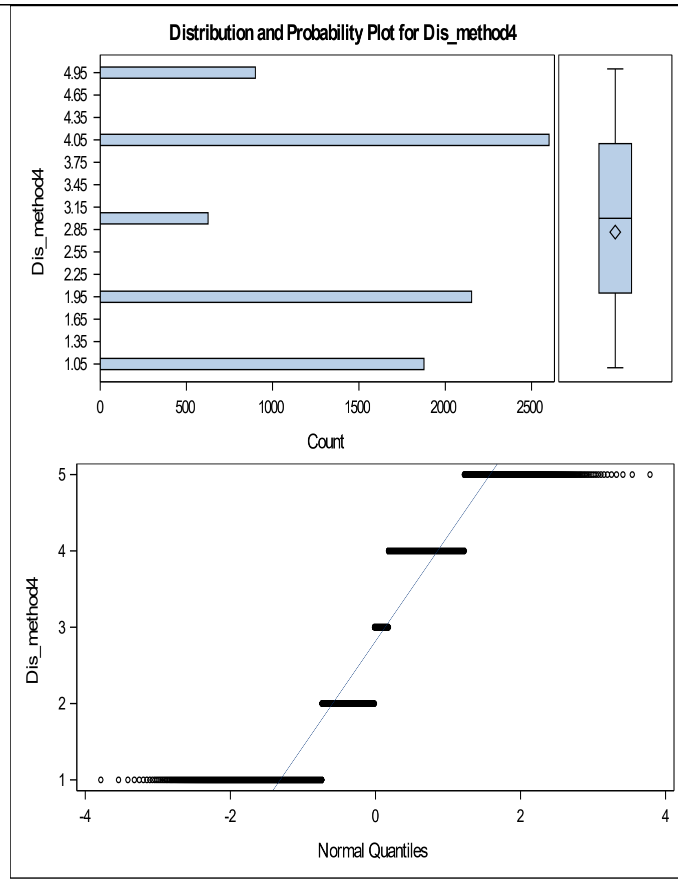


Table C3: Variable statistics - Household Head's education

Moments			
N	8156	Sum Weights	8156
Mean	1.8934	Sum Observations	15443
Std Deviation	1.0358	Variance	1.0730
Skewness	0.6091	Kurtosis	-1.107
Mean	1.8934	Std Deviation	1.0358
Median	1.0000	Variance	1.0730
Mode	1.0000	Cell (%)	3.0000

Frequency Counts			
1.None	4259	52.2	52.2
2.Primary	1104	13.5	65.8
3.Middle/JSS	2196	26.9	92.7
4.Secondary+	597	7.3	100.0

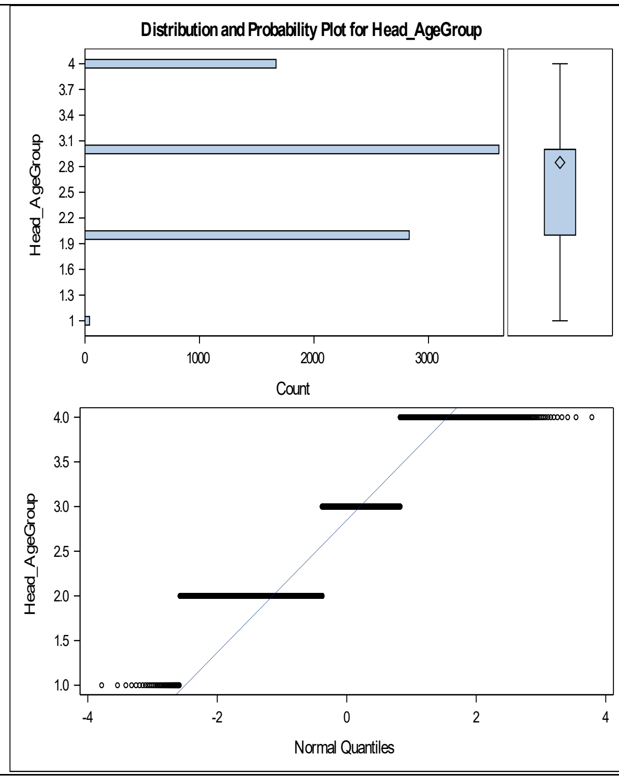


Table C4: Variable Statistics - Sex of household head

Moments			
N	8156	Sum Weights	8156
Mean	1.2410	Sum Observations	10122
Std Deviation	0.4277	Variance	0.1829
Skewness	1.2110	Kurtosis	-0.5334
Mean	1.2410	Std Deviation	0.4277
Median	1.0000	Variance	0.1829
Mode	1.0000		1.000
Frequency counts			
Value	Count	Cell (%)	Cum (%)
Male 1	6190	75.9	75.9
Female 2	1966	24.1	100.0

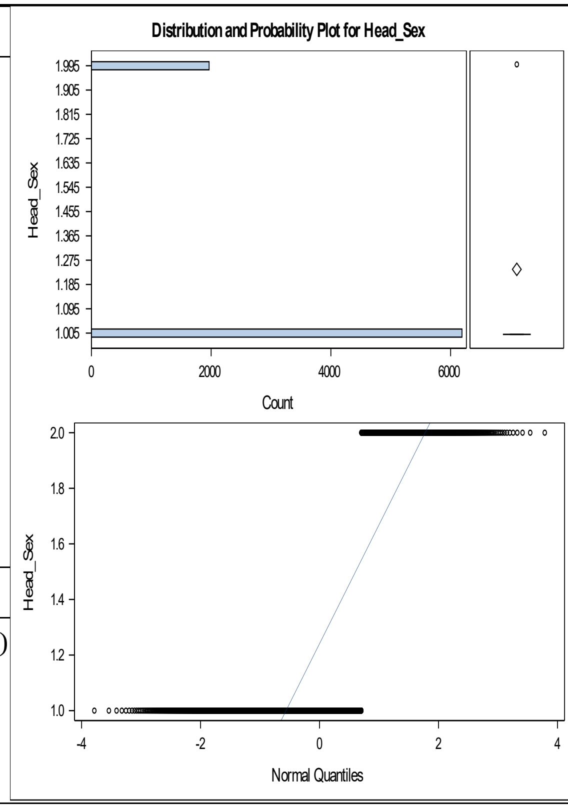


Table C5: Variable Statistics - Household head's ethnicity

Moments			
N	8156	Sum Weights	8156
Mean	17.1042	Sum Observations	139502
Std Deviation	8.9355	Variance	79.8422
Skewness	6.2711	Kurtosis	53.0377
Mean	17.1042	Std Deviation	8.9354
Median	15.0000	Variance	79.8422
Mode	21.0000	Cell (%)	85.0000
Frequency Counts			
Akan	2404	29.5	29.5
Ga/Dangme	343	4.2	33.7
Ewe	745	9.1	42.8
Guan	317	3.9	46.7
Gruma	540	6.6	53.3
Mole Dagbani	2981	36.5	89.9
Grusi	479	5.9	95.7
Mande	148	1.8	97.6
Non-Ghanaian	124	1.5	99.1
Others	75	0.9	100.0

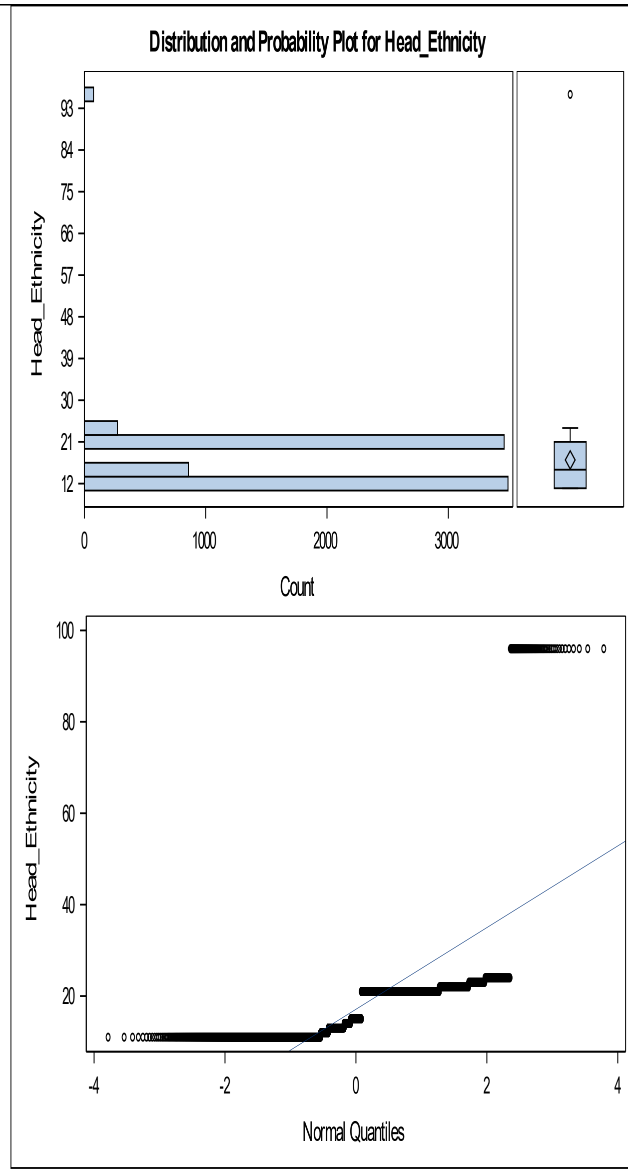


Table C6: Variable Statistics - Household wealth index

Moments

N	8156	Sum Weights	8156
Mean	2.3178	Sum Observations	18904
Std Deviation	1.3860	Variance	1.9210
Skewness	0.6525	Kurtosis	-0.9120
Mean	2.3178	Std Deviation	1.3860
Median	2.0000	Variance	1.9210
Mode	1.0000	Cell (%)	4.0000

Frequency Counts

Value	Count	Cell (%)	Cum (%)
1.Poorest	3336	40.9	40.9
2.Second	1656	20.3	61.2
3. Middle	1249	15.3	76.5
4.Fourth	1066	13.1	89.6
5.Richest	849	10.4	100.0

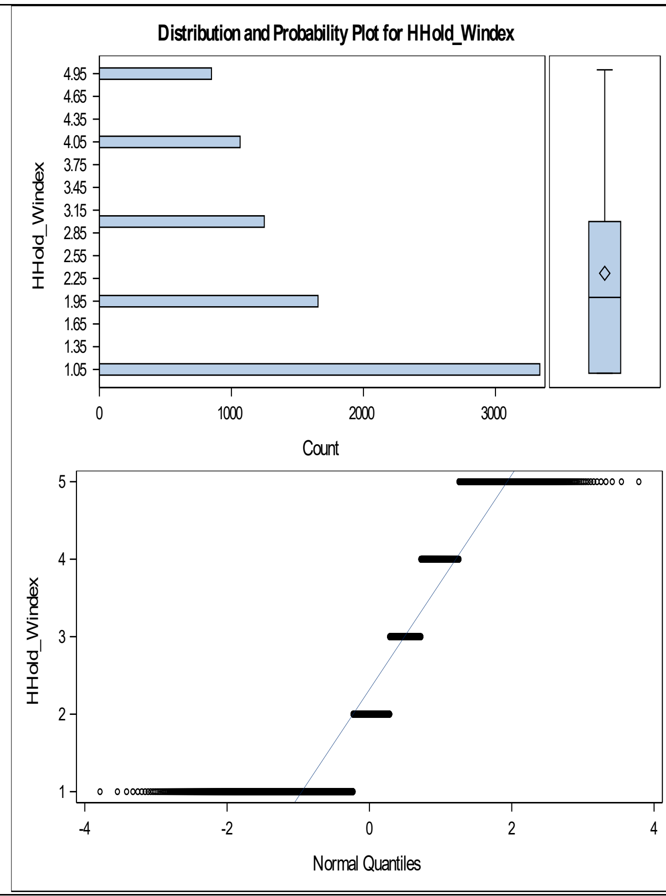


Table C7: Variable Statistics - Household head's religion

Moments			
N	8156	Sum Weights	8156
Mean	1.7290	Sum Observations	14102
Std Deviation	0.9265	Variance	0.8585
Skewness	1.0391	Kurtosis	-0.0094
Mean	1.7290	Std Deviation	0.9265
Median	1.0000	Variance	0.8585
Mode	1.0000	Cell (%)	3.0000
Frequency Counts			
Value	Count	Cell (%)	Cum (%)
1.Christian	4389	53.8	53.8
2.Moslem	2104	25.8	79.6
3.Traditional	1147	14.1	93.7
4.Other	516	6.3	100.0

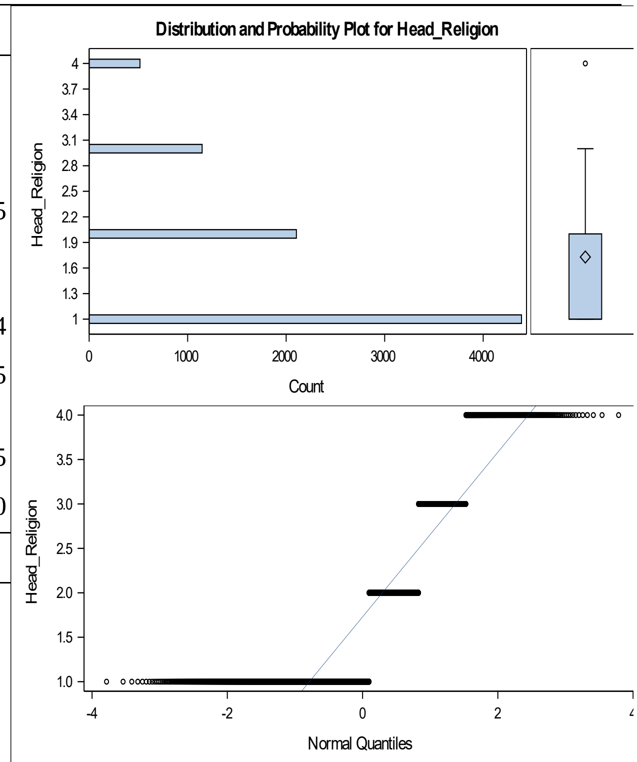


Table C8: Variable Statistics - Household head's age group

Moments

N	8156	Sum Weights	8156
Mean	2.8475	Sum Observations	23225
Std Deviation	0.7404	Variance	0.5483
Skewness	0.1783	Kurtosis	-1.0120
Mean	2.8475	Std Deviation	0.74049
Median	3.0000	Variance	0.5483
Mode	3.0000	Cell (%)	3.0000

Frequency Counts

Value	Count	Cell (%)	Cum (%)
1.15–21	40	0.5	0.5
2. 22–40	2832	34.7	35.2
3. 41–60	3615	44.3	79.5
4. 60+	1669	20.5	100.0

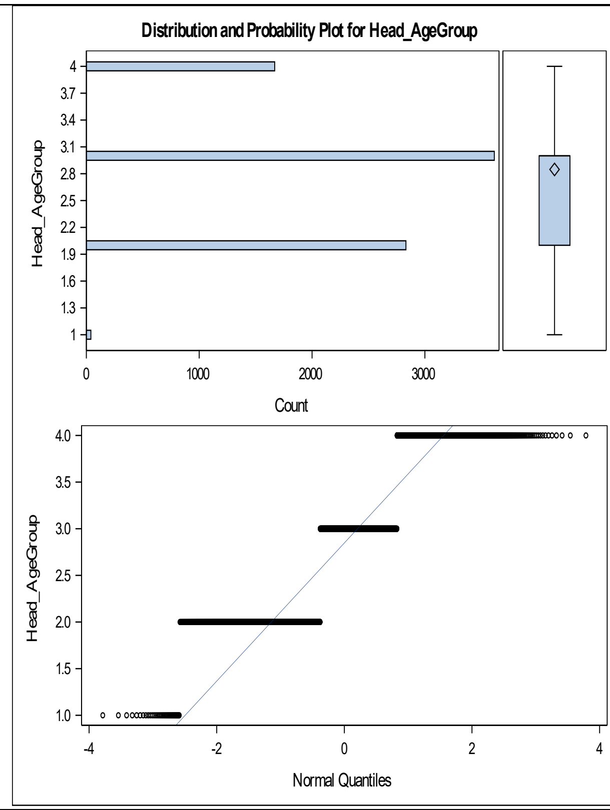
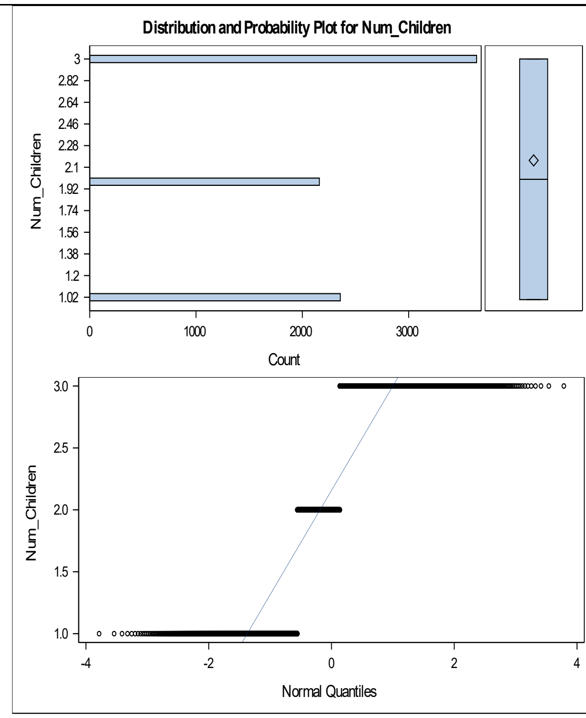


Table C9: Variable Statistics - Number of children in household

Moments			
N	8156	**Sum Weights**	8156
Mean	2.1571	**Sum Observations**	17594
Std Deviation	0.8429	**Variance**	0.7105
Skewness	-0.3035	**Kurtosis**	-1.5272
Mean	2.15718	**Std Deviation**	0.8429
Median	2.0000	**Variance**	0.7105
Mode	3.0000	**Cell (%)**	2.0000
Frequency Counts			
Value	Count	Cell (%)	Cum (%)
1 Child	2357	28.9	28.9
2 children	2160	26.5	55.4
3 Children	3639	44.6	100.0



APPENDIX D

Table D1: Variables, variable names, type of variable and values of the variables.

	Dependent Variable	Variable Name	Variable Type		Variable Values
	Response Variable 1				
	Is physical discipline appropriate for good child upbringing?		Categorical	1	'yes'
				2	'no'
	Response Variable 2 Most used discipline method by household	Disc_ Method	Categorical (nominal)	1	All methods (random use of any method)
				2	Non-physical and psychological method discipline methods
				3	Non-physical methods
				4	Physical discipline methods
				5	Psychological methods
	Predictor/Independent Variables				
1	Religion of Household Head	Head_Religion	Categorical (nominal)	1	Christian
				2	Moslem
				3	Traditionalist
				4	Other Religion
2	Age group of household head	Head_AgeGroup	Categorical (Ordinal)	1	15-21 years
				2	22-40 years
				3	41-60 years
				4	61+years
3	Number of children in household (2-14)	Num_children	Categorical (Ordinal)	1	1
				2	2
				3	3+
4	Sex of household head	HHSEX	Categorical (nominal)	1	Male
				2	Female

CONTINUATION OF TABLE D2

5	Wealth quintile	HHold_Windex	Categorical (Ordinal)	1	Poorest
				2	Second
				3	Middle
				4	Fourth
				5	Richest
6	Ethnicity of Household head	Ethnicity of head	Categorical (nominal)	1	Akan
				2	Ga/Dangme
				3	Ewe
				4	Guan
				5	Mole-Dagbani
				6	Other
7	Education of household head	Helevel	Categorical (Ordinal)	1	None
				2	Primary
				3	Middle/JSS
				4	Secondary+
8	Locality of Household	HH6	Categorical (nominal)	1	Urban
				2	Rural
9	Region of household	HH7	Categorical (nominal)	1	Western
				2	Central
				3	Greater Accra
				4	Volta
				5	Eastern
				6	Ashanti
				7	Brong Ahafo
				8	Northern
				9	Upper East
				10	Upper West