University of Nairobi

School of Mathematics

ISSN: 2410-1397

Master Project in Social Statistics

Modeling The Key Determinant of Child Labour In Kenya

Research Report in Mathematics, Number 0037, 2020

Joseph Magu Thiong'o

November 2020

Submitted to the School of Mathematics in partial fulfilment for a degree in Master of Science in Social Statistics

Modeling The Key Determinant of Child Labour In Kenya

Research Report in Mathematics, Number 0037, 2020

Joseph Magu Thiong'o

School of Mathematics
College of Biological and Physical sciences
Chiromo, off Riverside Drive
30197-00100 Nairobi, Kenya

Master Thesis
Submitted to the School of Mathematics in partial fulfilment for a degree in Master of Science in Social Statistics

Submitted to: The Graduate School, University of Nairobi, Kenya

Abstract

Child labour is an effect of many factors that are addressed in the MDGs, SDGs and verious policy documents. In the listen years , programmes and policies have not been established out to address the issues of child labour owing to the fact that this has not been adequately captured or analysed in national data and statistics.

The main objective of this study is to investigate the key determinants of child labour in Kenya. The study focused on children of the aged between 5 and 14 years using the KNBS Household survey Data of 2017. Mixed effect binary logistic regression was conducted to analyse the data. The explanatory variables are: child age and sex,household size,family head gender, type of household residence, relationship of a child to the household head, household head level of education, hours spent by a child on household chores, average monthly household income and expenditure and area of residence.

The model results show that the age of a child, the highest grade attended by the household head (household head education),average household monthly income, hours spent by a child in carrying out household chores and area of residence are important determinants of child labour in Kenya. The findings indicate that the chance for child to be engaged in work increases with age. Household income has negative influence on the chance for child labour. Higher level of education of the household head decreases the chance of sending child to work. In addition, increase in hours spent on household chores increases possibility of child labour. Lastly, the type and are of residence significantly affect child labour.

Policy interventions to be enhanced for reduction of child labour are improving households living conditions by increasing their average monthly income. Raise adult literacy levels. Reduce hours spent by children in taking household chores and enhance gender equality in education. Address regional disparities in probability of child labour by allocating more educational resources to the devolved government units with high child labour probability.

Master Thesis in Mathematics at the University of Nairobi, Kenya.
ISSN 2410-1397: Research Report in Mathematics

DISTRIBUTOR: School of Mathematics, University of Nairobi, Kenya

Declaration and Approval

I the undersigned declare that this dissertation is my original work and to the best of my knowledge, it has not been submitted in support of an award of a degree in any other university or institution of learning.

	25/11/2020
Signature	Date

JOSEPH MAGU THIONG'O
Reg No. I56/7700/2017

In my capacity as a supervisor of the candidate's dissertation, I certify that this dissertation has my approval for submission.

	25/11/2020
Signature	Date

Dr Ann Wang'ombe
School of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: awangombe3@gmail.com

Dedication

I dedicate this work to my beloved family members my mother Nancy Nyambura, wife Naomi Wanjiru, daughter Nancy Nyambura,my mother Nancy Nyambura, brothers; Paul Ngaii, John Macharia and George Karuma, Uncle [Kimani Jona], my friends Dr Magana, Dr Mugetha , Mr Lamec Mwimbanda, Peter Njuguna and Lucy Mugo for the great support they accorded me. Above all I thank God for the good health and strength he gave me through the entire period.

Contents

List of Tables

List of Figures

List of Acronyms

AIC	Akaike Information Criterion
BIC	Baysian Information Criterion
cloglog	Complementary Log log link function
DIC	Deviance Information Criterion
EC	European Commission
GER	Gross Enrolment Rate
GLM	Generalized Linear Models
GLMM	Generalized Linear Mixed Models
HH	Household
KNBS	Kenya National Bureau of Statistic
ILO	International Labour Office
LPM	Linear Probability Model
LRT	Likelyhood Ratio Test
SDGs	Sustainable Development Goals
MDGs	Millennium Development Goals
MELR	Mixed Effect Logistic Regression
MICS	Multiple Indicator Cluster Survey
MLE	Maximum Likelihood Estimates
MoEST	Ministry of Education Science and Technology
O.R	Odd Ratio
PRSP	Poverty Reduction Strategy Paper
SDGs	Sustainable Development Goals
UN	United Nations
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations International Children's Emergency Fund (United Nations Children's Fund)

Acknowledgments

First, I thank God for enabling me in different ways, Dr Ann Wang'ombe for accepting to be my supervisors. For Chiromo library department staff, for providing a conducive environment for my studies. More so, all my lecturers, thank you. Finally, all my friends Josphat Mwongera Paul,twin brother Paul Ngaii, James Njagi, Simon Karani, ,Benjamin Macharia, Benjamin Muriithi, John Macharia, Elius Kinyanjui, Eunice Mwangi, Rose Mwega, Alex Nyamari, Samuel Nyabuto, Ann Wambui Kiragu, Kenneth Oborah, Chrispinus Rakula, Grace Kinyua and Sarah Kendi. Your support and contribution.

Joseph Magu Thiong'o

UON Nairobi, 2020.

1 Introduction

2 INTRODUCTION

2.1 Introduction

By most indications, child labor appears to be a phenomenon of major proportions in the developing world. The International Labour Organization (ILO) estimates that around 190 million children between 5-14 years of age were economically active in 2004 (ILO 2006).1 This figure represents slightly less than 16 percent of all children in this age group.
Child labour phenomenon is considered a major obstacle to Millennium Development Goals (MDGs) and Poverty Reduction Strategy Paper (PRSP) (Global March Against Child Labour, 2013). The issue of Decent work and economic development form one of the Sustainable Development Goals (SDGs). The child labour problem is a major concern in Child development since it affects their ability to acquire quality education and healthy development. This study focuses on the key determinants of child labour in Kenya.

A study on Child labor across the developing world Patterns and correlations(Jean 2007) shows the extent of child labour in All regions as 12.4%, that of Sub-Saharan Africa 21.2% and Middle East and North Africa 6.0%. In Kenya households engage in different economic activities in order to generate income. Currently Kenya has forty seven counties brought about by devolution. This study will focus attention on key determinants of child labour in Kenya.

2.2 Back ground of study

There is no universally agreed upon or defined measure of child labor. Definitions of the relevant age of the child and the type of work vary across different studies and surveys. Three basic categories of child work are identified as economic activity, child labor, and hazardous work. Severe limitations characterize these indicators. They are hard to measure, exclude important activities that children undertake in own household (chores), and are subject to important seasonal variations. Data allows us to examine child economic activity and child labor. Child economic activity is defined as all paid work and certain forms of unpaid work (such as unpaid work in own household enterprises).

Child labour doesn't have a universally accepted definition. In most cases it is difficult to differentiate between the concept and definition of child labour. It has been argued by some authors that it is impossible to come up with a definition of child labour that captures all its features due to the complexity of child labour phenomenon. Weston (2005) equates child labour to social construct which differs by actors, history, context and purpose.

Different organizations dealing with issues of child labour differ in the concept and definition of child labour.The following are some concepts of child labour.

The ILO Concept and Definition of child labour

ILO draws its concept of child labour from the ILO Minimum Age Convention No. 138 of 1973. The convention sets the minimum age which a child should engage in any form of employment as 15 years. Any form of work that in violation of Convention No. 138 is considered child labour and illegal and should therefore be stopped. ILO went ahead and introduced child work and child labour. According to ILO child work may be acceptable while child labour should be eliminated.

Four groups of children engaged in work or labour are identified below:

- Working children
- Children who are economically active aged between 5 to 11 years are considered to be engaged in child labour. Also children aged between 12 to 14 years and are eco-

nomically active are considered to be engaged in child labour except if they engage in light work for less than 14 hours per week.

- Children in hazardous work. Hazardous work is any form of work likely to cause harm to the health, safety and moral development of a child. This group involves children working in mines, construction or other hazardous activities and includes children aged 18 years and below and working 43 hours or more per week.
- The last group is those children involved in the worst forms of labour as defined by ILO Convention No. 182. It includes children in forced or bonded labour, armed conflict, prostitution and pornography, and illicit activities.

The "worst forms of child labour" comprise: (a) slavery and forced labour, including child trafficking and forced recruitment for armed conflict; (b) the use of children in prostitution and pornography; (c) the use of children in illicit activities; and (d) any activity or work by children that, by its nature or conditions, is likely to harm or jeopardize their health, safety or morals – often referred to as "hazardous work" (ILO, 2013).

Two points come out from this view of ILO. Firstly, from the four groups we see the first group covering all activities which seem to right according to ILO, while the second and third groups cover activities of child labour which needs to be eliminated and the fourth group gives a picture that needs an urgent action for elimination.

ILO does not include children under the age of 5 since they are considered too young to be working. The second point is that this definition only considers work that can generate income such as production of goods and services. There is no mention of household chores such as cooking, cleaning or taking care of young ones. Gibbons, Huebler & Loaiza (2005) in criticizing the ILO definition, argue that it is too narrow since it underrates the harm that work has on children especially girls who mostly perform household work compared to boys.

The UNICEF Concept and Definition of Child Labour

The ILO definition of child labour has been expanded by UNICEF by considering the domestic work done by children apart from the economic work. Child labour is defined by UNICEF as follows:

- Children 5 -11 years engaged in any economic activity, or 28 hours or more domestic work per week.

- Children 12-14 years engaged in any economic activity (except light work for less than 14 hours per week), or 28 hours or more domestic work per week.
- Children 15-17 years engaged in any hazardous work.

The goodness with the UNICEF definition is that it captures all work done by children. This definition also gives an indication of child labour which is harmful to children's physical or mental development. However, it is of limited value for an analysis of the trade-off between work and school attendance.

Child labour, the MDGs and SDGs

Previously, Child labour was linked to MDGs through their cause and effect. The relationship between child labour and MDGs is that the problems that MDGs seek to address are what affect child labour processes. The major problem that affect child labour are adressed in MGD goal 1 that seek to eradicate extreme poverty and hunger. Lack of education , especially for girls, is illustrated by low levels of primary school enrolment. This is in MGD goal 1 and 2 whose aim is to achieve universal primary education and promote gender equality and empower women respectively (UN, 2015). Other MDG goals focus on combat HIV/AIDS , improve maternal health, reduce child mortality, malaria and other major diseases and ensure sustainability of the environmental. The MDG goal 8 fucused to develop a global partnership for development.This eradicate problem related to poor public policy. (UN, 2015) Futhermore, child labour makes achievement of MDGs difficult due to engaged of children in child labour. This is shown by national statistics , that show this children are left out of the programs and policies.The above goals were re-energised by the SDGs. The challenges related to child labour are addressed in SDG goals 1,2,4 & 5 which are, end poverty in all its forms everywhere, end hunger; achieve food security and improved nutrition and promote sustainable agriculture, ensure inclusive and equitable quality education and promote lifelong learning opportunities for all and achieve gender equality and empower all women and girls respectively. In goal 5 child marriage and child bearing among adolescents is also addressed. (UN, 2017)

World Bank, describe the observed destruction of child labour on long term investment as a serious threat to development (Weston, 2005).On the other hand, ILO views child labour from the point of view of the long run effect it has on children in their day to day participation in income generating activities in the household while UNICEF looks beyond investment and economic activity, and incorporates work done domestically and not to the interest of the child (Huebler, 2006). This not withstanding, there is need to agree on a universal definition of child labour for the purpose of policy making. The following subsection considers both ILO and UNICEF definitions.

It is clearly known that there are young people who work in the labour market for wages while others work in the family without pay. There has been a growing interest on the part of international organizations, researchers and governments to understand the factors that determine both supply and demand of child labour.

The 18th century ushered the Industrial Revolution in Great Britain. It is during this revolution that child labour which was a social problem associated with industrial production and capitalism, and accepted in agricultural societies in the early ages start to be opposed. The opposition of child labour became enormous in other countries that were industrializing in the following century (Shahrokhi, 1996).

History of Child labor across the developing world

This study state that child economic activity rates, on average, is roughly 1 in 5 children work, though there exists significant variation in child economic activity rates across countries. The regional breakdown reveals that, compared to most regions, AFR is unique; the region has the highest child economic activity rate with roughly 1 in 3 children working. The gender breakdown reveals that, in the vast majority of countries, boys are more likely to work than girls.
The exceptions to this general rule are predominately found in AFR. On average, roughly 1 in 4 boys and 1 in 5 girls work. Examining the relationship between the share of children working (i.e., the economic activity rate) and the share of children working and not attending school, the two appear to be strongly and positively Turning now to the distribution of economically active children, in terms of gender, in most countries, boys outnumber girls. On average, the ratio of working boys to working girls is 3 to 2.

Next, in terms of sector of activity, in virtually all countries, most working children are in agriculture, followed by services and then manufacturing. On average, out of every 10 working children, roughly 7 are in agriculture, 2 in services, and 1 in manufacturing. The gender breakdown reveals that these two results apply more or less equally to both genders. The regional breakdown reveals differences in the relative importance of the three sectors, but these differences were not testable owing to small sample sizes. In terms of child labor, on average, roughly 1 in 8 children are engaged in child labor, although there is substantial variation across the sample countries, most of it intra-regional rather than inter-regional. The patterns for child labor are largely consistent with the patterns for child economic activity. The child labor rate is highest in AFR, with roughly 1 in 5 children engaged in child labor. Across countries, the child labor rate for boys tends to be higher than for girls. Further, the mean child labor rate for boys is higher than for girls.

The development of any nation or institution depends on its human resource. Child labour hinders human resource development by denying children chance to be in learning institution. (JAMON, 2010). Child labour greatly contributes to the poverty rate among the community. This is because child labor increases the dropouts of children from schools; it also decreases the school enrollment rate. In contrast, education equips one with life skills which enables one to move from poverty to prosperity. Education is a part of any solution of reducing and eliminating the child labor. Despite the potential disadvantages and hardships for children engaged in child labour, school enrollment and attendance rates are evidence of child labour in some schools. It is not clear what explains the status of child labour rate in Kenya. This is because as very few studies have been conducted about the problem of child labour. In order to address the key determinants of child labour, there is a need to have a clear understanding of the nature and causes of child labour in Kenya. Without this knowledge, it would be difficult to formulate policies and interventions to reduce or/and eliminate the phenomenon of child labour across the country. This study examines the factors influencing the family's decision to subject her child to work . This is done with especial attention to gender, regions, and residence differences. The analysis of having or not having a child experience child labour is a binary response variable. Hence, Logistic model is used to analyse and estimate the determinants of child labour in Kenya.

2.3 Statement of the problem

In Kenya, several unpublished research works on child labour have been carried out. These study essentially tried to establish the link between factors that would make a child participate to the labour market. This approach of the analysis is abound in the literature of child labour but lacks of a consensual definition of child labour due to the complexity of the phenomenon. This study aims at bring additional light on child labour. The main objective is to identify the characteristics and determinants of child labour in Kenya. JAMON (2010) described the development of any nation or institution depends on its human resource.

Child labour hinders human resource development by denying children chance to fully concentrate in learning institution educational activities which equips one with life skills that enables one to move from poverty to prosperity. Child labour hence greatly contributes to the poverty rate among the community. This is because it may increases the school dropouts and decreases the school enrollment rate. In contrast, Education is a part of any solution of reducing and eliminating the poverty. Despite the potential disadvan-

tages and hardships for children engaged in child labour, school enrollment, attendance rates and performance are evidence of child labour in some schools.
It is not clear what explains the status of child labour rate in Kenya. This is because as very few studies have been conducted about the problem of child labour. In order to address the key determinants of child labour, there is a need to have a clear understanding of the nature and causes of child labour in Kenya. Without this knowledge, it would be difficult to formulate policies and interventions to reduce or/and eliminate the phenomenon of child labour across the country.
This study examines the factors influencing the family's decision to subject her child to work . This is done with especial attention to some fixed variables of child and household characteristics and random variables of the residence differences. The analysis of having or not having a child experience child labour is a binary response variable. Hence, mixed effect logistic regression model is used to analyse and estimate the determinants of child labour in Kenya.

2.4 Objectives of the study

The general objective
The main objective of this study is to model the key determinants of child labour in Kenya by employing Mathematical regression.

The specific objectives of the study include:

(i) to identify the major aspects of child work that contribute to child labour.

(ii) to estimating the supply equation of child labour using the mixed effect logistic model.

(iii) to identify policy implementation that would minimise child labour in Kenya.

2.5 Research Questions

This study attempts to answer the following questions:

(I) Which factors contribute significantly to child being involved in labour?

(II) How does various factors contribute to child labour?

(III) Are there policies to intervene on the issues of child labour by different stake holders?If so, which policies are recommended?

2.6 Justification of the study

The future of every nation lies in her youth who are its children. This can only be realized if the children are well equipped with the necessary skills to enable them take over from the aging population. Child labour from literature available indicates that it depends to a great extent on the income of the family and the educational level of parents concerned. This study is expected to send light into the "problem" of child labour in the region and especially in the study area. It will also bring awareness of the issues that lead to child labour to the local community and how to address them. The findings of the study will help in knowing the magnitude of the problem in the study area. If the recommendations are implemented they can help minimize the effects of the problem of child labour in the study area. The research findings will also add to the existing literature of knowledge. The research findings and recommendations will stimulate interest in the area and call for further research in future.

2.7 Scope of the study

Geographically, the study will cover three Counties among the 47 counties of Kenya (Kilifi, Kitui and Busia). There are disparities in terms of political stability and population density among the Counties.

The study focus on the determinants of child labour in Kenya for children aged 5 to 14 years. The study uses household survey data comprises on all the three zones. (KNBS 2017)

2.8 Organization of the study

This research has been organized into five chapters. The first chapter introduces the research, identifies the key problem under investigation and states the specific objectives for the research.

It further, asks the relevant research questions, gives a justification for the topic, and defines its scope. The second chapter contain a review of relevant literature on child labour and the framework of relevant variables. Chapter Three contains the data source, explains the methods and procedures used in the study, and defines the key data variables. The results will be represented in chapter four. This is a very important chapter in the research because it provides the information to answer the research questions raised. The findings will be based on the data analyzed in this chapter. Chapter five contains conclusion and recommendations for policy implications.

3 LITERATURE REVIEW

3.1 Introduction

In this chapter, we review literature on how other researchers have applied determinants of child labour in analyzing similar data as used in this study. The section on literature review focuses on the objectives of the studies, the methods of data collection, data analysis techniques, the variables used in the study, and results of the data analysis. The other section examines the conceptual framework of the study. The chapter ends with a summary of the reviewed literature.

3.2 Literature review

Chaubey(2007) identified and examined the relationship of child labour with a large set of possible factors with the data for analysis on 175 countries. Regression model was conducted using total child labour (the ratio of children employed in their total population) as dependent variable. The explanatory factors used in the study were: female literacy in the country (mother's education), economic growth rate (GDP growth rate) and the proportion of population with income less than a dollar a day (poverty). The results found, showed that a 1 percentage point rise in female literacy can reduce total child labour in a country by 25 percent and that of female child by 30 percent. Similarly, a 1 per cent point reduction in population below one dollar a day can reduce total child labour by 21 percent. Economic growth is quite an effective factor in reducing female child labour. Girls are more vulnerable to economic downturns than boys. Child labour is a major cause for low enrollment in secondary school. This difference would hence suggest the need for better targeting of girl child for economic support during seasonal or catastrophic poverty.

Rubkwan (2008) investigated the various factors that influence a household's decision of sending a child to work. This presented a detailed empirical analysis of the determinants of child labour in Thailand. Econometric analysis is carried out using data from the Thailand labour force survey (National Statistic Office Thailand, 2003). Multiple regression model was estimated using number of hours children worked in the last 7 days before the survey (child time of work) as the response variable. The explanatory variables in this paper were classified as: The 'children' characteristics that are age and wage. The 'household' characteristics which are the household's monthly income, region of residence, number of children, gender of household head, age of household head, parental education, and occupation of household head. However, the 'school' and 'community' characteristics were not incorporated due to the limitation of the dataset. The estimates of the model indicated that wage impacted significantly on the time that children allocate to work. Age had significant effect on boys but insignificant on girls. This implied that the older the boy became, the more time he would be allocated to work. Boys and girls in urban areas were found to work fewer hours compared to their counterparts in the rural areas. The effect was more on boys than in girls. The size of the household affected the working time of children positively. This implies that households that had more members had their children working more hours. Girls tend to benefit from household's head age. This implies the more the age, the more the time boys are allocated work compared to girls. Educated parents were found to allocate fewer working hours for their children. The occupation of the household head was also found to affect the working time of children. Children from households in which occupation of the household head is related to agriculture were found to be involved in some form of work.

Laurent(2010) did a study on characteristics and determinants of child labour in Cameroon using data from the Cameroonian survey on employment and informal sector. In this study, Binary probit and Tobit models were estimated using child time of work as the response variable. Predictor variables used were grouped into child's, household and household head's characteristics. The child's characteristics include the sex, age, relationship with the household. The household characteristics that include the income, residence, size and the composition of household. Finally, household head's characteristics were the level of education, the type of employment, the age and sex. They found in the estimated models, an increase in income together with increase in household size resulted to a reduction in the time that a child spent on work. Also their study found that an increase in the place of residence together with an increase in the level of education of household head could result to a reduction in the time spent by children working. All the variables in the study were found to be statistically significant except the level of education of the household head which was statistically insignificant. This implies that an increase in adults' income by 1 per cent would result to a decrease in time spent by children to work by more than 2 hours. The reduction tends to favor girls since their working time reduces by 3 hours compared to boys' 1 hour 30 minutes. Time of work is more sensitive in urban areas compared to rural areas in regards to household income variations. Additionally, when the household size increases by one person, the study found that it results in reduction of time spent by children working by 0.6 per cent in favor of boys (-0.64%) than girls (-0.55%). In terms of residential areas, children in urban areas tend to benefit more compared to their counterparts in the rural areas. Age of a child also has a significant effect on the time spent on work by a child. Older children allocate more time to economic work. Being a child of the household head reduces time in hours that young boys are allocated to work in economic activities by 0.33%. Also children living with their parents in the urban areas were found to allocate 0.69% less time for work. Households headed by females were found to allocate less time to work for their children compared with those headed by males. In addition households headed by females allocated more time domestic works so as to enable their children spend time studying. These households have higher and significant probability of sending children to school.

Moyi (2011) examined the causes and magnitude of child labour in Kenya. The data used for this study was drawn from the second round (2000) of the Multiple Indicator Cluster Survey (MICS). MICS is a household survey program that UNICEF developed to assist member states with collecting data to monitor the condition of children and women. These data are used to assess progress towards the goals set at the 1990 World Summit for Children at two points, mid-decade and end-decade. The first round of MICS (mid-decade) was conducted in 1995/1996 and the second round (end-decade) of surveys was conducted in 2000. A third round of MICS, conducted in 2005 to 2006, is used to monitor progress towards the MDGs. The data used in this study was drawn from the second round of the MICS. Multinomial logistic regression was used in this paper. The variables used in this study were classified as to the children, household, and community characteristics that influence child labour and school participation. Age, gender, and the relationship with the household head and the number of young children; siblings between 0 and 3 years, gender of the head of household, and education of the head of household as years below or above eight years, have impact on school and/or work participation. The income of the household, and the type and place of residence are some household characteristics that may impact school and/or work participation.
This study hypothesized that the socioeconomic status and structure of the household would have a strong effect on child labour as well as many children who were working to be attending school. The study found children's activities to be affected by their age and gender, how they are related to the household head, education level of the household head, wealth of the household and the children present in the household. Although the study found that children of the household head were less likely to be working only and attending neither school nor work; however, they had a higher probability of combining both school and work. Urban children were found to combine school and work four times less likely compared to their rural counterparts. The study recommended that policy makers formulate policies which would factor education inequality dimension between children who combine work and school and those who do not combine if the effect of working is going to hinder children from attaining education.

Tifow (2014) did a study to investigate the key determinants of child labour in Somalia using Somalia 2006 Multiple Indicator Cluster Survey (MICS) data. He used binary logistic regression model where the response variable was either child work or child schools. The explanatory variables were child's age and gender, parental education, family size, father and mother income, household regions, and area of residence. The model results show that; The child whose mother has primary education are 20.5 % less likely to be engaged in work than those whose mothers have no education .Those of secondary education, tertiary and non curriculum education had 5.4 % 24.4% and 10.1% respectively less likely to be involved in work than those whose mother had no education. The child whose father has primary education are 19.7 % less likely to be engaged in work than those whose mothers have no education .Those of secondary education, tertiary and non curriculum education had 1.4%, 8% and 1.5% respectively less likely to be involved in work than those whose mother had no education. A unit increase in the household wealthy, makes a family 0.702 times (30 %) more likely to withdraw their children from the work compared to the children from poor households. A child that live in urban areas is 27 % less likely to work than rural children. An increase in household size by one person would decrease the child probability of being engaged in work by 0.3%. An year increase in child age increases child labour by 20.5% . Boys are found to be 39% less likely to work relative to their girls counterparts.

Satriawan(2018) investigated the nature of relationship between parental income and child labor supply in Indonesia.This study benefited by panel data from the last two waves of Indonesia Family Life Survey (2007 and 2014). In this study a linear and quadratic regression model model are used. The response variable for the regression model were the child labour hours while the predictor variable were the squared natural log of both parents' income. Different models were run; these include the full sample model, boys sample model and girls sample model, urban sample model and rural sample model. Other variables not included in the model include child & household characteristics that include sex and age of child, child still in school, father and mother's education in years, the number of male, female adults in the household and the number of older and younger children in the household.From the results child labor is more prevalent in rural areas than in urban areas for all age group, while higher age is associated with an increase in child labor incidence. Fathers have a higher average income than mothers implying they are the breadwinner. From the regression, as wages increases, child labor hours decreases, and at high level of wages, father's income and child labor are substitutes.From the quadratic model the relationship between father's income and child's working hours is persistent strong and significant in rural areas.Age is not found to be significant.Parental education initially appears to be important determinant that lower child labor. However, after controlling for household/parental fixed-effects, the significant effects disappear.Having a sibling in the household decreases estimated child labor hours, which may indicate that the burden of working is likely to be shared between children.

3.3 Conceptual framework

The action of household decision to either send or not to send a child to work is usually a function of a number of factors. Child labour has several adverse effects on human capital development. The researcher will investigated the factors that influence child labour using quantitative methods.
The frequently used variables in previous literature review about child labour can be classified into 4 groups, namely; the children, the household, the school, and the community characteristics.

The Children characteristics include; gender, age and relationship with the household.

The Household characteristics include; parental education, occupation of household head, gender of household head and the household size.

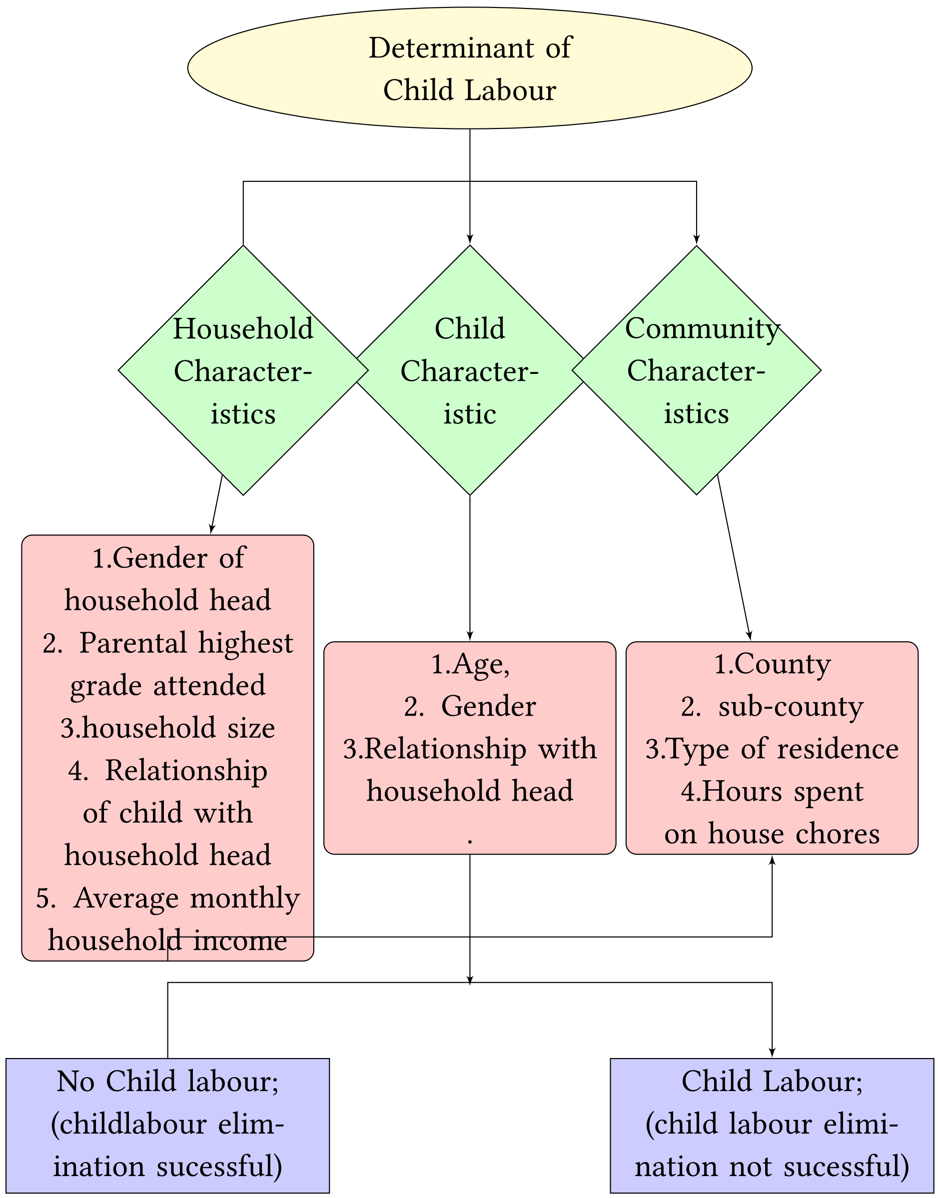
The community characteristics include; county, sub-county, type of residence(urban, rural or periurban) type of school and quality of school attended by child.

In summary, successful interventions of these factors may reduce and eliminate the child labour adjasting for other factors. **Figure 2.1: Conceptual Framework**

3.4 Summary of the Literature Review

From the reviewed literature, the issue of child labour is evident. The challenge would be, lack of reliable data capturing the magnitude of the problem. Factors that influencing child labour include; single parenting, ignorance of the parents, poverty and strong socio-cultural beliefs. Poverty has come out as a major determinant of child labour in developing countries, With the majority of the children engaged in agricultural activities. A common socio-cultural belief common among rural dwellers is that child labour is part of a training program, and therefore considers them as essential contributors to household incomes. Laws are important in fighting this vice, but if no enforcement is patchy the situation is not helped - this is the case (as observed in the literature) beyond the formal sector where children get involved in agricultural and domestic chores. For the purpose of this study, child labour is defined as any activity, economic or non-economic, performed by a child, that is either too dangerous or hazardous and/or for which the child is too small to perform and that has the potential to negatively affect his/her health, education, moral and normal development. The legal definition of a child in Somalia is anyone who has not reached the age of maturity, which is 18 years (CSFR, 2012). It is accepted that children under 5 years are not physically capable of undertaking work of any significance, whether economic or non-economic. The target group for the study, therefore, comprised all children aged 5 to 14 years, engaged in economic or non-economic activities (including housekeeping/household chores in their own parent'/guardians' household).

Figure 1. Conceptual Framework



4 MIXED EFFECT BINARY LOGISTIC REGRESSION

4.1 Introduction

This chapter discusses the methodology employed in the study. The methods and the procedures discussed have been motivated by the nature of the problem in the research in the previous chapter and the type of data available.

4.2 Theoretical Model

Binary Logistic Regression Model

When modelling dependent variable with two possible categorical outcomes Binary Logistic Regression (Binary Logit Model) is commonly used. Instead of being continuous these variables have only two possible outcomes that are nominal. The model is used to predict which outcome is likely to occur. For the purpose of analysis, the possible outcome are assigned two dummy codes (1 and 0). These dummy variables are used are used since the probability of an event occurring must lie between 0 and 1. For linear regression model the outcome are continuous and my lie to values less than 1 or greater than 1, this explain why this model cannot model binary categorical outcome. The logistic regression model is a type of generalized linear model that extends the linear regression model by linking the range of real numbers to the 0-1 range.
The binary response variable can be defined as;

$$f(x) = \begin{cases} 1, & if \quad there \quad is \quad child \quad labour, \\ 0 & if \quad there \quad is \quad no \quad child \quad labour. \end{cases} \tag{1}$$

Here, $P_r(Y=1) = \pi$ and $P_r(Y=0) = 1-\pi$ the $E(Y) = \pi$. Y is known as a Bernoulli random variable.
If the binomial variable Y with parameters n trials and probability of child labour in each trial being π . The mean probability of child labour is $E(Y) = n\pi$ and variance $\sigma^2 = n\pi(1-\pi)$, then the probability density function is;

$$f(y) = \binom{n}{y}\pi^y(1-\pi)^{(n-y)}, \quad \text{where} \quad y = 1,2,3,....,n$$
$$= \binom{n}{y}(\frac{\pi}{1-\pi})^y(1-\pi)^n \tag{2}$$
$$= exp \quad [ylog(\frac{\pi}{1-\pi}) + nlog(1-\pi) + log\binom{n}{y}]$$

This equation is in the form of exponential family of distributions which is indicated in the appendix with $\theta = log \quad (\frac{\pi}{1-\pi})$ and hence $\pi = (\frac{e^\theta}{1+e^\theta})$, $c \quad \oslash \quad = 1$, $b(\theta) = nlog \quad (1+e^\theta), c \quad (y, \oslash) = \quad log \quad (\binom{n}{y})$.
Hence,

$$\begin{aligned} b(\theta) &= \text{nlog} \quad (1+e^\theta) \\ b'(\theta) &= \frac{ne^\theta}{1+e^\theta)} = n\pi, \quad and \\ b''(\theta) &= \frac{ne^\theta}{(1+e^\theta)^2} = n\pi(1-\pi) \end{aligned} \tag{3}$$

From this ,
$E(Y) = n\pi(1-\pi)$ and $var(Y) = n\pi(1-\pi)$ as expected.
Like ordinary regression, Logistic regression can have one or multiple explanatory variables. These explanatory variables may be continuous measurement or categorical variables. The categorical variables may be converted to continuous variables or rather be expressed as dummy variable.
When more than one predictor variables are used to explain the response variable, the model is expressed as,

$$logit(\pi_i) = ln(\frac{\pi_i}{1-\pi_i}) = \sum \quad x_{ij} \quad \beta_i \tag{4}$$

i = 1,2,3,....,n
j = 0,1,2,3, ...,n
n = number of predictor varables
β_i = *regression parameter of variable i.*

x_{ij} is a matrix of the predictor variables.
θ is vector of the unknown regression parameters.

In the model a fixed change in x_i may have less impact when probability of Y occurring (π_i) is near 0 or 1 than when (π_i)is near the middle of its range. In this model the response and explanatory variables are linearly related since the range of the probability (0-1) is extended to real numbers by transformation of link function logit.
Expressing the equation 3.4 in exponential form and solving for π_i,

$$
\begin{aligned}
& ln(\frac{\pi_i}{1-\pi_i}) = \sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i \\
& (\frac{\pi_i}{1-\pi_i}) = e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i)) \\
& \pi_i = (1-\pi_i) \quad (e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i) \\
& \pi_i = 1(e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i) - \pi_i(e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i) \\
& \pi_i + \pi_i(e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i) = e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i \\
& \pi_i(1+e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i) = e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i \\
& \pi_i = \frac{e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i)}{1+e^{(}\sum_{j=0}^{j=p} \quad x_{ij} \quad \beta_i)}
\end{aligned}
\tag{5}
$$

This function give a Logistic curve.

Generalised Linear Mixed Model

Generalized linear mixed models (GLMMs) are an extension of linear mixed models to allow response variables from different distributions, such as binary and count responses. Alternatively, they are extension of generalized linear models such as logistic regression to include both fixed and random effects of the predictors. Hence mixed models. The general form of the model in matrix notation) is:

$$y = X\beta + Z\mu + \varepsilon$$

Where **y**is a $N \times 1$ column vector,the outcome variable **X** is a $N \times p$ matrix of the p predictor variables; β is a $p \times 1$ column vector of the fixed-effects regression coefficients. **Z** is the $N \times q$ design matrix for the **q** random effects complement to the fixed **X**. **u** is the $q \times 1$ vector of the random effects the random complement to the fixed and ε is $N \times 1$ column vector of the residuals, the part of **y** that is not explained by the model $X\beta + Z\mu +$.
The different between LMMs and GLMMs is that the response variables can come from different distributions besides Gaussian normal distribution .In addition, rather than modeling the responses directly, some link function is often applied, such as a log link. Let the linear predictor, η, be the combination of the fixed and random effects excluding the residuals.

$$\eta = X\beta + Z\mu$$

The generic link function called $g(.)$ that relates the outcome **y** to the linear predictor η we get

$$\begin{aligned} \eta &= X\beta + Z\mu \\ g(.) &= link \quad function \\ h(.) &= g^{-1}(.) = inverse \quad link \quad function \end{aligned} \tag{6}$$

So the model for the conditional expectation of **y** is:

$$g(E(y)) = \eta$$

it is conditional because it is the expected value depending on the level of the predictors. We could also model the expectation of **y**:

$$E(y) = h(\eta) = \mu$$

with **y** itself equal to :

$$y = h(\eta) + \varepsilon$$

For a binary outcome, we use a logistic link function and the probability density function, or PDF, for the logistic. These are:

$$
\begin{aligned}
& g(.) = \eta = X\beta + Z\mu \\
& \eta = log_e \quad \frac{\mu}{1-\mu} \quad making \quad \mu \quad the \quad subject \\
& e^{\eta} = \frac{\mu}{1-\mu}, \quad e^{\eta}(1-\mu) = \mu, \quad e^{\eta} - \mu(e^{eta} = \mu, \\
& e^{\eta} = \mu + \mu e^{\eta}, \quad \mu(1+e^{\eta}) = e^{eta} \\
& \mu = \frac{e^{\eta}}{1+e^{\eta}} \\
& h(.) = \mu = \frac{e^{\eta}}{1+e^{\eta}} \\
& P.D.F = \frac{e^{\frac{x-\mu}{s}}}{s(1+e^{-(\frac{x-\mu}{s})})^2} \\
& E(X) = \mu \\
& Var(X) = \frac{\pi^2}{3}
\end{aligned}
\tag{7}
$$

4.3 Model parameters interpretation

In the model the predictor variable x_i include either be a continuous measurement or a dummy variable corresponding to a categorical variable. In the the logistic regression model both the continuous and/or categorical predictor variables can be included. β_i (where i= 0, 1, 2, ...) are unknown regression coefficients that need to be estimated.
the interpretation of the predictor parameter depend on the type of predictor variable. whether it is a continuous or categorical variable

4.3.1 Continuous predictor variable.

Consider, the x_i continuous predictor variable in the model taking k and $(k+1)$ values (an increase of one unit) . The odds associated to a given predictor is;
odds for predictor $x_i = \frac{\pi_i}{1-\pi_i} = e^{(\Sigma_{j=0}^{j=i})} = e^{\beta_0+\beta_1 x_1+\beta_2 x_2 ...,\beta_i x_i}$

odds for predictor $(x_i = k)$=$\dfrac{\pi_i}{1-\pi_i} = e^{(\Sigma_{j=0}^{j=k})} = e^{\beta_0+\beta_1 x_2+\beta_2 x_2 \ldots,+\beta_i k}$

odds for predictor $(x_i = k+1)$= $\dfrac{\pi_i}{1-\pi_i} = e^{(\Sigma_{j=0}^{j=k})} = e^{\beta_0+\beta_1 x_1+\beta_2 x_2 \ldots,+\beta_i (k+1)}$

odd ratio for one unit increase of x_i from k to k +1 ,

$$odd \quad ratio = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2 \ldots,+\beta_i (k+1)}}{e^{\beta_0+\beta_1 x_1+\beta_2 x_2 \ldots,+\beta_i k}} = \frac{e^{\beta_i (k+1)}}{e^{\beta_i^k}} = e^{\beta_i}$$

$O.R = e^{\beta_i}$

Therefore , the exponential of the estimated parameter term (e^{β_i}) is the change in likelihood of variable x_i occurring for every unit additional measure of the predictor variable. The effect of explanatory varible x_i to the odd of response occurring is explain in for different ways according to its value. this explanations are;

1. If $\beta_1 = 0$ then $O.R = 1$. This imply the predictor variable is not significant in predicting the response event. If $O.R = 1$ the event occurring for both predictor and response variable is the same.
   Thus the odds ratio of 1 is used as a reference point for interpretation of other odds ratio.

2. If $\beta_i < 0$, then $0 < O.R < 1$. The response event is e^{β_i} times less likely to occur for every unit increase in predictor variable.Alternatively, the effect can be stated that, the response event is $100(1 - O.R)$ % less likely to occur for every unit increase in predictor variable.

3. If $1 < O.R < 2$ the effect can be stated that, the response event is $100(O.R - 1)$ % more likely to occur for every unit increase in predictor variable.

4. If $\beta_i > 2$, then $O.R => 2$. The event is $O.R$ times more likely to occur for every unit increase in predictor.

4.3.2 Categorical predictor variable.

When using categorical variables it is important to use dummy variables for its different levels. To do this, first select one level of the variable as a reference group and then create a dummy variable. For dichotomous dummy variable;

1. The group that posses characteristic of interest is assigned the level of $X = 1$.
2. The reference group (otherwise) is assigned level of $X = 0$.

For instance, the predictor dichotomous categorical variable for type of residence may include urban or rural; Here the reference group may be picked as being rural labeled others and assigned a level of $X = 0$ while urban would be assigned a value of $X = 1$. These dummy may be illustrated in case form as;

$$Residence(X) = \begin{cases} 1, \quad if \quad individual \quad is \quad from \quad urban, \\ \\ 0 \quad Otherwise \quad (if \quad individual \quad is \quad not \quad from \quad urban). \end{cases} \tag{8}$$

The odd relating to the dummy variables for variable x_i are;

odds for predictor x_i and dummy X = O,

$$Odds \quad when \quad (X=1) = \frac{\pi_i}{1-\pi_i} = e^{(\Sigma_{j=0}^{j=k})} = e^{\beta_0+\beta_1 x_2+\beta_2 x_2 ...,+\beta_i(1)}$$

$$Odds \quad when \quad (X=0) = \frac{\pi_i}{1-\pi_i} = e^{(\Sigma_{j=0}^{j=k})} = e^{\beta_0+\beta_1 x_2+\beta_2 x_2 ...,+\beta_i(0)}$$

Odd ratio for level 1 and 0 of predictor x_i

$$odd \quad ratio = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2 ...,+\beta_i(1)}}{e^{\beta_0+\beta_1 x_1+\beta_2 x_2 ...,+\beta_i 0}} = \frac{e^{\beta_i(1)}}{1} = e^{\beta_i}$$

$$O.R = e^{\beta_i}$$

Therefore , the exponential of the estimated parameter term (e^{β_i}) is the change in likelihood of response variable in relation to variable x_i occurring for group associated to dummy $X = 1$ as compaired to the reference group, controlling/adjusting for the other variables. These odd rations are explain in for different ways according to its value. This include;

1. If $\beta_1 = 0$ then $O.R = e^{\beta_i} = 1$. This imply the levels in the predictor variable is not significant in predicting the response event adjusting for all the other variables. If $O.R = 1$ the event occurring for both groups (the compared and reference) are the same.
   Thus the odds ratio of 1 is used as a reference point for interpretation of other odds ratio adjusting for all the other variables.

2. If $\beta_i < 0$, then $O.R = e^{\beta_i} < 1$. The response event is e^{β_i} times less likely to occur in a given group compared to the reference group. Alternatively, the effect can be stated that, the response event is $100(1 - e^{\beta_i})$ % less likely to occur in a group compaired to the reference group adjusting for all the other variables.

3. If $1 < O.R(e^{\beta_i}) < 2$ the effect can be stated that, the response event is $100(e^{\beta_i} - 1)$ % more likely to occur in a given group compared to the reference group adjusting for all the other variables.

4. If $\beta_i > 2$, then $O.R = e^{\beta_i} > 2$. The event is e^{β_i} times more likely to occur in stated group group compared to the reference group adjusting for all the other variables.

4.4 Statistical Inference Testing For Mixed Models Parameters

Statistical inference refers to the process of drawing conclusions from the model estimation. It involve data analysis to deduce properties of an underlying probability distribution. Inferential statistical analysis infers properties of a population, for example by testing hypotheses and deriving estimates. It is assumed that the observed data set is sampled from a larger population.

Inferential statistics can be contrasted with descriptive statistics. Descriptive statistics is solely concerned with properties of the observed data, and it does not rest on the assumption that the data come from a larger population.
Unlike OLS and GLM parameters which asymptotically converge to known distribution, mixed models parameters do not have nice asymptotic distributions to test against(Wisconsin,2019). This complicates the inferences made from mixed models.These complexity

include a penalty factor (shrinkage) which is applied to the random effects in the calculation of the likelihood (or restricted likelihood) function the model is optimized to.This results in distributions which are no longer chi squared or F. This penalty factor also complicates determining the degrees of freedom to associate with the estimate of a random effect of the model. Another source of complications is in testing the significance of a variance parameter. Since the $\sigma^2 \geq 0$, a test of zero is on the border of the parameter space. Tests of parameters are valid only on the interior of their space and not on the border. The correlation structure within the data complicates using bootstrap procedures to test these statistics which do not have known distributions. Parametric bootstraps which can more easily account for the correlation in the model are more typically used for inference in mixed models than bootstraps, which are non-parametric. In the process we compute the GLM, the parameters β_i values for each regressor and inference about the parameters. This enables us test some hypothesis and goodness of model fit.

Inference about fixed effects parameters

Fixed effects tests are typically done with either Wald or likelihood ratio tests (LRT). Wald and LRT tests are equivalent with the assumptions of asymptotic distributions and independent predictors. If the data set size is not large enough to be a good approximation of the asymptotic distribution or there is some correlation amongst the predictors, the Wald and LRT test results can vary considerably. The LRT is generally preferred over Wald tests of fixed effects in mixed models. For linear mixed models with little correlation among predictors, a Wald test using the approach of Kenward and Rogers (1997) will be quite similar to LRT test results. The most reliable inferences for mixed models are done with Markov Chain Monte Carlo (MCMC) and parametric bootstrap tests. Some common test for coefficients are available for mixed models include; **The Wald test.** Tests of the effect size which is scaled using the estimated standard error. **The LRT (Likelihood Ratio Test.)** Tests the difference in two nested models using the Chi square distribution. **The Profiled confidence interval**. While not a formal test, an interval which does not contain zero indicates the parameter is significant. **The Parametric bootstrap.** Assumes the model which restricts a parameter to zero (null model) is the true distribution and generates an empirical distribution of the difference in the two models. The observed coefficient is tested against the generated empirical distribution. Since the distributions of coefficients are only approximately asymptotical, two or more of the above tests are generally done to confirm results of tests that are inconclusive.

The inference about Model Parameters $\widehat{\beta_i}$ made using the Wald test statistic is $z = \frac{\widehat{\beta_i}}{SE}$. Z has an approximate standard normal distribution when $\beta_i = 0$ and z^2 a chi-squared distribution with $df = 1$. When the sample size is less than 30 we use the t-test.

Parameter hypothesis testing:

For H_0: $\beta = 0$ and $H_1 : \beta \neq o$, we use Wald's test statistic $z^2 = (\frac{\widehat{\beta_i}}{SE})^2 \sim \chi_1^2$ and reject null hypothesis if calculated z score is greater than the tabulated one. Alternatively, we use a P-value test statistic where $P - value = 2 \times prob(Z > z)$. Here, we reject the null hypothesis if the p-value is less than level of significance.
For H_0: $OR = 1$ and $H_1 : OR \neq 1$ the test statistic is a CI. We reject H_O at α level of significance if the $100(1-\alpha)\%$ CI does not contain 1.

Confidence Interval

The $100(1-\alpha)\%$ confidence interval for the model parameter for β_i is
$\beta_i \pm z_{\frac{\alpha}{2}} SE(\beta_i)$ where SE is the standard error of β_i.

Inference about Random parameters

It include testing if the variance parameter is equal to zero (0).The test which are in common use for the variance parameter include; **LRT (Likelihood Ratio Test)** The variance parameter of a generalized mixed models does not have a known asymptotic distribution.The LRT for these variance parameters at times can be poor estimates. We recommend treating these p-values with caution.The LRT test of a variance parameter equalling zero will be conservative ,larger p-value, due to the test being on a boundary condition ($\sigma^2 \geq 0$). If the p-value is small enough to be significant the finding is likely good.In some areas twice the LRT p-value is used as a formal test.(SSCC,2019)

Goodness of model fit

Goodness of model fit is a test for overall significance of the model. For mixed models this the overall significance is done for fixed effect and random effect separately. It is a test statistic for the hypothesis that all parameters that are in the fitted model are not significantly different from that of saturated model. The test is made using the deviance of GLM. The deviance of a GLM is defined by,

$$deviance = -2 \times log \quad (\frac{L_M}{L_S}) = -2 \times logL_M + 2 \times logL_S \tag{9}$$

The, L_S denote the maximized log-likelihood value for the saturated model, model with all the predictors. This is the model that provides a perfect fit to the data.It has additional parameters, its maximised log likely hood is atleast as large as the maximised log likely hood of fitted model. Saturated model is the most complex model possible from each of the observations.

L_M denote the maximized log-likelihood value for the fitted model.For the mixed effect model we use the nested model; model with one less fixed predictor for overall significance of fixed effect and one less random factor to test overall significance of random effect. If the model has only one random factor then the fixed effect model become a random. The details of this are illustrated in appendix III.

The deviance is the log-likelihood ratio statistic for comparing fitted model to the saturated model. A test statistic for the hypothesis that all parameters that are in the saturated model but not in fitted model equal zero. GLM software provides the AIC , BIC, log-likelyhood the deviance for both models. The, chisquaree value and p-values are provided for the fitted model. For some GLMs, the deviance D has approximately a chi-squared distribution with degree of freedom the difference between number of observations and parameters.

$$D \sim \chi^2 \quad (n-p)$$

Where,
n = number of observations.
p = number of parameters is fitted model.
D = The deviance statistic that describe a lack of fit in fitted model. The larger the value of D the poorer the fitted model. A fit is significant if the test give a significant result or

the p-value is larger than the significant level.

Alternatively, in testing for the adequacy of the fitted model one can also use the null model (a model with intercept only) and fitted model.

In this case, the null and alternative hypotheses are:
H_o: The null model is better fit.
H_1: The fitted model is a better fit.
The deviance statistic describes goodness of fit. The significant fit is one where the test gives a significant result.

4.5 Model Specification and mis-specification

These study has assume sequential decision making process. The household head make the first decision for the child to be at work. It and will use binary logistic model. Simultaneous decision making process is one in which the choice is made by the household head from a pool of decisions on whether to sens child to school or work.(Moyi,2011)The two options are contrasted in this study. The household face two discrete options in which they try to maximize utility. It choose from the two mutually exclusive events of either having **child work** or **child does not work**.

Using Child does not work as the reference group for this model. The probability of a child engaging in work is given by;

$$Prob[Y_i = 1 \mid i = 1, 2, 3, ...] = \pi_i = \frac{e^{(\Sigma_{j=0}^{j=i} X_{ij}\beta_i)}}{1+e^{(\Sigma_{j=0}^{j=i} X_{ij}\beta_i)}}$$

Where, $\beta = (\beta_1, \beta_2, \beta_3, ..., \beta_p)^T$ and $X_{ij} = (X_{i1}, X_{i2}, X_{i2}X_{i3}, ..., X_{ip})$.

In this study the explanatory variable in vector X will be categorized into individual characteristics, household characteristics and community characteristics. **The individual characteristics** include the age and sex of the child. **The household characteristics** include the size of the household (the number of people in the household), the sex and the highest grade ever attained by the family head, the household monthly and net monthly income. **The community characteristics** include the type of residence (urban, rural or peri urban), the type of dwelling structure (permanent or non permanent)

4.6 Estimation Method

Engagement of children in work is a dichotomous variable. Binary model is hence an appropriate choice for this study. The dependent variable takes the values of child involved in labor or child not involved in labor. The objective of this study is to estimate the relationship between the probability of a child labor given the individual, household and community characteristics.

Linear probability model (LPM), probit model and logit model, are the estimations approaches that are employed when dealing with binary dependent variables (Damisa,2017 ; Gujarati, 2007). The LPM model violates the normality assumption. Also, its probability estimates usually lie outside the 0-1 bound. Since S-shaped probability model is needed for the study LPM is hence not a better choice. Probit or logit model are often chosen. this is because they not only correct the problem on non-normality but also their response variable tend to lie within the 0-1 range (Gujarati, 2007).

Categorical response regression are hence modeled using the logit, probit, loglog and the complementary loglog (Cloglog) link functions under symmetric and asymmetric assumptions. In binary regression logit is considered the default link. However; one may also choose probit or Complementary loglog links. Among the considered reasons include; the theoretical Considerations, influences by disciplinary traditions and characteristics of the data. The economists favour probit models while the toxicologist favour logit models. The complementary log log works best with extremely skewed distributions.(Damisa,2017)

The logit and probit are symmetric link functions where the Probit outperforming the logit link function. The loglog and cloglog are assymmetric (for skewed data) link functions where cloglog outperform the loglog link function. These link functions transforms probabilities to z-scores from the standard normal distribution. . When several of this models fit well, parameter estimates in probit models have smaller magnitude than those in logistic regression models. This is because their link functions transform probabilities to scores from standard versions of the normal and logistic distribution, but those two distributions have different spread. The standard normal distribution has a mean of 0 and standard deviation of 1. The standard logistic distribution has a mean of 0 and standard deviation of 1.8. When both models fit well, parameter estimates in logistic regression models are approximately 1.8 times those in probit models. The parameters of the model will be estimated using a MLE technique. The likelihood function for the model is specified as;

$$L(\frac{\beta}{y}) = \Pi_{i=1}^{n} \quad exp\{y_i log(\frac{\pi_i}{1-\pi_i}) \quad + \quad nlog(1-\pi_i) \quad + \quad log(\binom{n}{y})\} \tag{10}$$

The logarithm is a monotonic function and similarly, the maximum of the likelihood function will also be a maximum of the log likelihood function and vice versa. Thus, taking the natural log of Eq. 3.9 yields the log likelihood function:

$$L(\beta) = \Sigma_{i=1}^{n}\{y_i log(\frac{\pi_i}{1-\pi_i}) \quad + \quad nlog(1-\pi_i) \quad + \quad log(\binom{n}{y})\} \tag{11}$$

The maximum (log) likelihood estimates are the values of that maximize the log likelihood function. The critical points of a function that represent the maxima and minima occur when the first derivative equals to zero(0). If the second derivative, evaluated at that point, is less than zero, then the critical point is a maximum. Thus, finding the maximum likelihood estimates requires the first and second derivatives of the likelihood function. Given $log\frac{\pi_i}{1-\pi_i} = \Sigma_{j=0}^{p} x_{ij}\beta_j$ and after solving for π_i we get $\pi_i = \frac{e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}}{1+e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}}$ Using this , the log lokelyhood equation is simplified to;

$$\begin{aligned} L(\beta) &= \Sigma_{i=1}^{n}\{y_i \Sigma_{j=0}^{p} x_{ij}\beta_j + \quad nlog(1 - \frac{e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}}{1+e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}}) \quad + \quad log(\tbinom{n}{y})\} \\ L(\beta) &= \Sigma_{i=1}^{n}\{y_i \Sigma_{j=0}^{p} x_{ij}\beta_j + \quad nlog(1+e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}) \quad + \quad log(\tbinom{n}{y})\} \end{aligned} \tag{12}$$

To find the maximum points of the log likelihood function, set the first derivative with respect to each β_i equal to zero. Thus,

$$\begin{aligned} \frac{\partial l(\beta)}{\partial \beta_j} &= \Sigma_{i=1}^{n}\{y_i x_{ij} - \quad n \cdot \frac{1}{1+e^{\Sigma_{j=0}^{p} \quad x_{ij}\beta_j}} \quad \cdot e^{\Sigma_{j=0}^{p} \quad x_{ij}\beta_j} \quad \cdot \quad \frac{\partial \beta}{\partial \beta_j}(\Sigma_{j=0}^{p} x_{ij}\beta_j) \\ \frac{\partial l(\beta)}{\partial \beta_j} &= \Sigma_{i=1}^{n}\{y_i x_{ij} - \quad n x_{ij} \cdot \frac{e^{\Sigma_{j=0}^{p} \quad x_{ij}\beta_j}}{1+e^{\Sigma_{j=0}^{p} \quad x_{ij}\beta_j}}\} \\ \frac{\partial l(\beta)}{\partial \beta_j} &= \Sigma_{i=1}^{n}(y_i x_{ij} - \quad n x_{ij}\pi_i) \end{aligned} \tag{13}$$

The maximum likelihood estimates for β can be found by setting each of the $p+1$ equations in Equation 3.12 equal to zero and determining the value of each β_j. The second derivative is used to determine whether the critical points of the function are maximum or minimum. If a function f has a critical point for which $f^{'}(x)=0$ and the second derivative is positive at this point, then f has a local minimum here. If, however, the function has a critical point for which $f^{'}(x)=0$ and the second derivative is negative at this point, $f^{''}(x)<0$, then f has local maximum here. This technique is called Second Derivative Test for Local Extrema. (Lynn,1990)

$$
\begin{aligned}
\frac{\partial l(\beta)}{\partial \beta_j} &= \Sigma_{i=1}^{n}\{y_i x_{ij} - \quad n x_{ij} \cdot \frac{e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}}{1+e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}}\} \\
\frac{\partial^2 l(\beta)}{\partial \beta_j} &= \frac{\partial}{\partial \beta_j}\Sigma_{i=1}^{n}\{y_i x_{ij} - n x_{ij} \cdot \frac{e^{\Sigma_{i=1}^{p} x_{ij}\beta_j}}{1+e^{\Sigma_{i=1}^{p} x_{ij}\beta_j}}\} \\
\frac{\partial^2 l(\beta)}{\partial \beta_j} &= -\Sigma_{i=1}^{p} n x_{ij}\{\frac{x_{ij}\cdot e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}(1+e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}) - x_{ij}\cdot e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}\cdot e^{\Sigma_{j=0}^{p}} x_{ij}\beta_j}{(1+e^{\Sigma_{j=0}^{p} x_{ij}\beta_j})^2}\} \\
\frac{\partial^2 l(\beta)}{\partial \beta_j} &= -\Sigma_{i=1}^{p} n x_{ij}\{\frac{x_{ij}\beta_j \cdot e^{\Sigma_{j=0}^{p}} x_{ij}\beta_j}{(1 \quad + \quad e^{\Sigma_{j=0}^{p}} x_{ij}\beta_j)^2}\} \\
\frac{\partial^2 l(\beta)}{\partial \beta_j} &= -\Sigma_{i=1}^{p} n x_{ij}\{\frac{e^{\Sigma_{j=0}^{p}} x_{ij}\beta_j}{(1 \quad + \quad e^{\Sigma_{j=0}^{p}} x_{ij}\beta_j)} \cdot \frac{1}{(1 \quad + \quad e^{\Sigma_{j=0}^{p}} x_{ij}\beta_j)} \cdot x_{ij}\} \\
\frac{\partial^2 l(\beta)}{\partial \beta_j} &= -\Sigma_{i=1}^{p} n x_{ij}\pi_i(1-\pi_i)x_{ij}^{T}
\end{aligned}
\qquad (14)
$$

This equation (Eq. 3.13) is the variance covariance matrix of the parameter estimates. The other equations (3.12 and (3.13) require numerical iterative techniques for their solution. Since they are nonlinear equations, the most popular method for solving systems of nonlinear equations in practice is the Newton- Raphson method (Scott). The MLE has the desirable statistical properties of normality, efficiency and consistency asymptotically (Long, 1997). However, mixed model might violate this properties.The Likelihood Ratio Test (LRT) of fixed effects is required. The LRT of mixed models is only approximately χ^2 distributed. For tests of fixed effects the p-values generated by the R soft ware are used to reject or fail to reject the null hypothesis. (SSCC 2019)

Table 1. Description of Variables Used in the Study

Variable	Definition
Dependant Variable	-
StatusOCLabour	status of child labour
Explanatory Variable	
Child Characteristic	
Age	Age of child in year
Sex	The male or female Gender
RelationHhhGuardian	Relationship of a child with household head
Household Characteristics	
fhead	Family head gender; Male or Female
hhsize	Household size; refer to number of people residing in a given hous
HGradeAttended	Highest Grade Attained by parent In scaled data from 1 for prepri
HHMonthlyIncome	Average Household Monthly Income in thousands of Ksh
HHMonthlyExpenditure	Average Household Monthly Expenditure in thousands of ksh
Hours.spent.on.Household.chores	hours spent by a child in household chores
Community Characteristics	
District	The county of residence; Busia, Kilifi, or Kitui
DIVISION	Specific Division in the county, include Kilifi,kitui and Busia
Tresidence	Type of household residence that include urban, rural and periurb

4.7 Variable Definition

4.8 Data Source

The data for the research is be obtained from household survey conducted by KNBS in 2017.These data will be used in this study to extract and analyse the key determinants of child labour in Kenya.

The sample was selected in four stages:

1. Three clusters (County) were selected in from the 47 counties in the whole nation. These are; Busia, kilifi and kitui.
2. five divisions were selected in each County .
3. Clusters were selected within the Location
4. Finally, households were randomly selected.

The questionnaires were administered to captured information on the household. This was done from parent, guardians and the eligible children. Among other factors the questions assess factors influencing households to send their children into work. A child is considered to be involved in child labour activities at the moment of the survey if during the week preceding the survey: A child age 5-11 year had at least one hour of economic work or had atleast 28 hours of domestic work per week. A child age 11-15 year had at least two hour of economic work or had atleast 28 hours of domestic work per week . Among other

The questionnaire had a number of questions addressing child labour . such questions include(UNICEF, 2007);
1. During the past one week did (name of the child)...

- miss school last week to; to work for pay, to work as unpaid worker in family business/farm, help at home with household chores and/or to learn a job.
- Worked for atleast 1 hr.
- Run or do any kind of business.
- Do any work for pay.
- Do any construction/ major repairs on household property.

2. Why did (name of the child) drop out of the school;to learn a job, to work for pay, to work as unpaid worker in family business/farm and / or help at home with household chores.
3. During the past one week what was the childs main activity; to Supplement family income, help pay family debt and/or help in household enterprise.

5 Data Analysis and Results

5.1 Introduction

This chapter presents the results of the study. The first section presents descriptive statistics of the variables included in the study. The second section presents the logistic regression results, and the last is discussion of the results. The R codes used in the analysis are in appendix III.

5.2 Descriptive Analysis

From the selected data set, there are a total of 7704 children between the ages of 5 to 14 years old . Out of these, 3841 children were involved in child labour during the last 7 days before the survey. This is 49.9 % of the observed children population in this area. This indication child labour situation in Kenya is serious. In this section the characteristics of the sample used in the study are presented. It focused on the effects of Child, household and community characteristics on child labour. The unit of analysis is a boys or a girls of child labourer aged 5 to 14 years.

Table 4.1 and table 4.2 shows some of the characteristics of children who work and those who do not work by looking at sample proportions in the categorical characteristics and sample means in continuous characteristics respectively. The percentage of children from Busia, Kilifi and Kitui are 34.8%, 32.1% and 33.1% respectively. In Busia the proportion of children under child labour (14.2%) is smaller compared to children not involved in child labour (20.5%). Similarly, Kilifi proportion of children under child labour (13.3%) is smaller compared to children not involved in child labour (18.8%). Kitui has the highest proportion of children under child labour (22.4%) compared to children not involved in child labour (10.7%).

This is also seen in Central, Chuluni, Jaribu, Matinyani, Mutitu, Mutongoni and Mwitika divisions. Central division has the highest number and proportion of child labour (443 which represent 5.8%) followed by Chuluni (338 which represent 4.4%)and Butula (324 which represent 4.2%) respectively. At 5% level of significance we conclude that the district of the child can be used to explain the situation of child labour.

Table 2. Descriptive Statistic by Working Children

Variable	Level	childlabour	non child labour	Total	chisquare	df
District	Busia	1096(14.2%)	1583(20.5%)	1583(34.8%)		
	Kilifi	1022(13.3%)	1452(18.8%)	2474(32.0%)	477.21	2
	Kitui	1723(22.4%)	828(10.7%)	2551(33.1%)		
DIVISION	BAHARI	301(3.9%)	461(6.0%)	762(9.9%)		
	BAMBA	110(1.4%)	178(2.3%)	288(3.7%)		
	BUTULA	324(4.2%)	540(7.0%)	864(11.2%)		
	CENTRA	443(5.8%)	304(3.9%)	747(9.7%)		
	CHONYI	105(1.4%)	155(2.0%)	260(3.4%)		
	CHULUNI	338(4.4%)	161(2.1%)	499(6.5%)		
	GANZE	43(0.6%)	69(0.9%)	112(1.5%)		
	JARIBU	59(0.8%)	57(0.7%)	116(1.5%)	566	17
	KIKAMB	218(3.6%)	373(4.8%)	654(8.5%)		
	MATAYO	248(3.2%)	343(4.5%)	591(7.7%)		
	MATINYANI	124(1.6%)	48(0.6%)	172(2.2%)		
	MUTITU	145(1.9%)	64(0.8%)	209(2.7%)		
	MUTONG	278(3.6%)	137(1.8%)	415(5.4%)		
	MWITIKA	148(1.9%)	5.(0.6%)	198(2.6%)		
	NA	350(4.5%)	118(1.5%)	468(6.1%)		
	NAMBALE	327(4.2%)	452(5.9%)	779(10.1%)		
	TOWNSHIP	138(1.8%)	239(3.1%)	377(4.9%)		
	VITENGENI	79(1.0%)	114(1.5%)	193(2.5%)		
Tresidence	PeriUrban	406(5.3%)	321(4.2%)	727(9.4%)		
	Rural	3071(39.9%)	3063(39.8%)	6134(79.6%)	25.57	2
	Urban	364(4.7%)	479(6.2%)	843(10.9%)		
fhead	Female	1199(15.6%)	1127(14.6%)	2326(30.2%)	3.809	1
	Male	2642(34.3%)	2736(35.5%)	5378(69.8%)		
RelationHhh	BioChild	3151(40.9%)	3149(40.9%)	6300(81.8%)	0.348	1
	Guardian	690(9.0%)	714(9.3%)	1404(18.2%)		
Sex	Female	1868(24.2%)	1883(24.2%)	3751948.7%)	0.01	1
	Male	1973(25.6%)	1980(25.7%)	3953(51.3%)		
Total	TOTAL	3841(49.9%)	3863(50.1%)	7704(100.0%)		

source: Computed From 2017 KNBS household baseline survey in Kenya

The highest number of children observed stay in rural area (6134 which is 79.6%) Those who stay in periurban and urban areas are 727 (9.4%) and 843(10.9%)respectively. Children staying in rural area are more likely to be involved in child labour.Rural area has the highest proportion of children under child labour (39.9%) compared to children not involved in child labour (39.8%).In periurban area the proportion of children under child labour (5.3%) is smaller compared to children not involved in child labour (4.2%). The urban area has the smallest proportion of children under child labour (4.7%) compared to urban children not involved in child labour (6.2%).Using the chisquare df= 2, the p value = 0.0000. Hence, Child labour is dependence of the type of residence at 0% level of significance.

The children from families that are headed by male are 69.78% while those headed by female are 30.22%. For families headed by female there are more cases of children involved in child labour (15.6% compared to their counterpart not involved in child labour(14.6%). In the families headed by male there are few cases of children involved in child labour (34.3% compared to their counterpart not involved in child labour(35.5%). Using the chisquare df= 1, the p value = 0.027. we therefor conclude Child labour is dependence of the type of residence at 5% level of significance.

A high proportion of children (81.8%) stay with their biological parents. 18.2 % stay with grand parents, brother,sister, friend, uncles or other people's families. There is no statistical significance to show that child labour can be explained by child relationship with the household head. The proportion of boys and girl observed are 51.3% and 48.7% respectively. Similarly, there is no statistical significance to show that child labour can be explained by the gender of the child.

5.3 The Mixed Effect Binary Logistic Results

This study tested whether the fitted model adequately describes the data available of child labour. The model has two parts, that of fixed effect and of the random effect. In this study we test the significant of the two parts separately. The null and alternative hypotheses for the study are: H_0 : null model is better fit ($\beta_1 = \beta_2 = \cdots = \beta_{11} = 0$) and H_1 : the fitted model is a better fit. ($\beta_1 \neq 0, \quad \beta_2 \neq 0 \quad \cdots \quad \beta_{13} \neq 0$). Table 4.5 show the value of the χ^2 for significance of fixed effect based on nested model1 and the saturated model. The value of the χ^2 is 152.07 with degree of freedom 1. This gives a $p-value = 0.0000$ We hence reject the null hypothesis at 0% level of significance. Thus, the fitted model adequately describes the data. . Table 4.6 show the value of the deviance for significance of random effect based on nested model2 and the saturated model. The value of the computed *loglik* is -337.31 with degree of freedom 14. This gives a $p-value = 1$. We hence reject the null hypothesis at 0% level of significance and conclude the fitted model adequately describes the data.

Table 4.4 reports the MELR results to estimate the determinants of child labour in Kenya. The dependent variable is categorical with two levels that are child labour or no child labour. The table presents the estimated coefficient for fixed effect predictor variables. In order to interpret the quantitative implications of the results we have an addition column with computed odds ratio.

Table 3. Radom Effect Results

Random effects							
Groups	Name	Variance	Std.Dev.				
Division	(Intercept)	0	0				
District	(Intercept)	0.1794	0.4236				
Number of obs:	3554	Groups:	Division	17	District	3	

Source: Computed from the 2017 KNBS household baseline survey

Table 4. Fixed effects Results

Fixed effects:	Estimate	Std. Error	z value	O.R	Pr(>|z|)	
(Intercept)	-2.090	0.304	-6.878	0.124	0.0000	***
Age	0.112	0.009	11.870	1.119	0.0000	***
SexMale	-0.045	0.053	-0.840	0.956	0.4011	
hhsize	0.003	0.011	0.296	1.003	0.7676	
fheadMale	0.061	0.060	1.022	1.063	0.3069	
TresidenceRural	0.061	0.114	0.539	1.063	0.5897	
TresidenceUrban	0.144	0.134	1.077	1.155	0.2815	
RelationHhhGuardian	0.051	0.070	0.732	1.053	0.4640	
HGradeAttended	-0.025	0.007	3.621	0.975	0.0003	***
Hours.spent.on.Household.chores	-0.017	0.003	-5.288	0.983	0.0000	***
HHMonthlyIncome	-0.006	0.003	-2.077	0.994	0.0378	*
HHMonthlyExpenditure	0.005	0.004	1.088	1.005	0.2766	
—						
Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’	0.1 ‘ ’	

Source: Computed from the 2017 KNBS household baseline survey

The reference group for sex is;	Female
The reference group for fhead is;	Female
The reference group for Tresidence is;	peri urban
The reference group for Relationship with hh is;	Biological Child

tabular

Table 5. Correlation of Fixed Effects:

Correlation of Fixed Effects:

(In	Age	SexMale	hhsize	fheadMale	TresidncRrl	TresdncUrbn	RltnHhhGrdn	HGradAttndd	Hrs.spn..H.	HHMnthlyInc	HHMnthlyE
Age	-0.27										
SexMale	-0.10	-0.07									
hhsize	-0.22	-0.06	0.05								
fheadMale	-0.11	0.08	-0.04	-0.22							
TresidncRrl	-0.36	0.03	0.02	-0.02	0.02						
TresdncUrbn	-0.34	0.04	0.02	0.10	0.03	0.79					
RltnHhhGrdn	-0.10	0.03	-0.03	-0.05	0.21	0.07	0.10				
HGradAttndd	0.00	-0.40	0.05	0.01	-0.02	0.02	0.03	-0.01			
Hrs.spn..H.	-0.10	-0.23	0.12	0.06	-0.02	0.00	-0.02	0.02	0.02		
HHMnthlyInc	-0.04	-0.04	0.03	0.02	-0.12	0.11	-0.09	0.00	-0.03	-0.01	
HHMnthlyExp	0.02	0.00	-0.02	-0.16	0.05	-0.05	-0.02	0.03	-0.03	0.01	-0.52
convergence	code:	0									

Source: Computed from the 2017 KNBS household baseline survey

Table 6. Test For Fixed Effect

	Df	AIC	BIC	logLik	deviance	Chisq	Chi df	Pr(>Chisq)	Sign
Nested model1 (fit11)	13	4524.2	4604.5	-2249.1	4498				
Fitted model(fit1)	14	4374.1	4460.6	-2173.1	4346	152	1	0.0000	***
Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' '				

Source: Computed from the 2017 KNBS household baseline survey

Table 7. Test For Radom Effect

	log lik	df	pChisq
fit1(fitted model)	-2173.07	14	
fit21(Nested model2)	-2341.07	13	
Deviance(calculated)	-337.31	14	1

Source: Computed from the 2017 KNBS household baseline survey

5.4 Discussion of the MELR

These study found different important findings. On individual characteristics, the study showed that, for every one year increase in a child age there is increases in child labour by 11.9%. this is statistically significant at 0.0000% level. This indicates that older children are more likely to be involved in child labour. This may be due to the fact that they are more physically mature and can take on more tasks and possibly fetch more higher pay. These results are similar to other studies such as Rubkan(2008), Laurent (2010), Moyi (2011)and Tifow(2014).

The male (boys) are 4.4% more likely to be involved in child labour than their counterpart girls. However, this is not significant at 0.05% level.

On household characteristics, for every one year increase in the highest grade attended by the household head, the house child odd of being involved in child labour decrease by 2.5%. this is statistically significant at 0.0000% level.This implies that children from highly educated family head are less involved in child labour. This is similar to other studies such as Rubkan(2008), Tifow(2014) and Satriawan (2018).

Household monthly income is another household characteristic. For every one thousand Kenya shilling increase in average monthly income, the possibility of the household child being engaged in child labour decrease by 0.6%. This is statistically significant at 0.05 % level. This implies that an increase in average income of a household decreases the possibility of sending children to work.Poor families hence are more likely to send their children to work. This is because poor families depends on earning from child employment to meet their food and essential needs. These results are similar to other studies such as Marcus(1998),ILO(2006), Chaubey(2007), Rubkan(2008), Laurent (2010), Moyi(2011),World Bank (2012), Tifow(2014) and Satriawan (2018).

Hours spent on household chores is also important determinant of child labour. For every one hour increase in household chores assigned to a child, the possibility of the household child being engaged in child labour increase by 1.7%. This is statistically significant at 0.0000 % level. this result is similar to that of Satriawan (2018).

In this study , the other predictors such as household size, gender of family head, type of residence and relation ship with household head effect are not statistically significant.

This study also sought to know whether children from different counties and locations in Kenya had higher probability of child labour than the other regions. These areas were obtained by clustered samples. In this study they are hence taken as random effect variable. The random effect is found to be significant with Deviance of 1. Children from different regions will hence have different probability of being involved in child labour.

6 Conclusion and Recommendation

This study examines key factors and characteristics determining child labour in Kenya using household survey data.
the chapter will contain;

6.1 Summary and Conclusion

This study examines key factors and characteristics determining child labour in Kenya using KNBS data – 2017 household Survey. To attain national goals of education for Kenya, knowledge of child labour and its impact is critical, yet complex hindrances make it difficult to define and examine. The UNICEF proposed expanded definition of child labour is important to this study because most working children participate in various tasks and household chores (Reynolds, 1991).

Children of the age between 5 to 14 years were sampled. Descriptive statistics were applied to the characterize child labour. A mixed effect binary logistic model fitted to estimate the determinants of child labour. The result showed that a child's age and gender, household head education and gender, average monthly household income, the area and type of residence are some of the factors influencing child labour in Kenya. The model indicated higher child labour rates were among older children as compared to the younger once.

Increase in average monthly income of a household decreases the chance of a child being engaged in child labour.

Parental education levels also indicated significance in that those children whose parents had higher number of years for grade attended were less likely to be involved in child labour than their counterpart from household with head having lower grades attained. Therefore, parental education affects the probability that the household child will be involved in child labour.

Taking the Consideration for area and type of residence the descriptive statistic indicate that children in rural areas were more likely to be child labourers than their counterparts from urban and peri-urban areas respectively.

6.2 Policy Recommendations

From the results, policies to be pursued to reduce child labour are those that:

- Improve households living conditions by increasing their average monthly income.
- Raise adult literacy levels by strengthening the existing programmes and establishing new ones in the whole republic.
- Reduce hours spent by children in taking household chores and enhance gender equality in education.
- Address regional disparities in probability of child labour by allocating more educational resources to the devolved government units with high child labour probability.

6.3 Area of Further Research

A census would be required to provide more information in order to arrive at a more complete list of determinants of child labour in Kenya. Moreover, more factors such as school and community characteristics should the added. Such factors should include; school accessibility,distance from home to school, type of school, school neighborhood and infrastructure,child's performance, child school attendance, father and mother education level and monthly income, number of younger and older sibling in the household should also be incorporated into the model since they may contribute to the unexplained effects. So far, little effort has been applied in investigating these factors.

Bibliography

[Alan, A. (2007)] Alan, A. (2007) *An Introduction to Categorical Data Analysis*, Second Edition, Florida ,John Wiley & Sons Inc.

[Chaubey (2007)] Jay Chaubey, J. , Perisic, m., Perrault, N., Laryea-Adjei Noreen , g., Khan, N. 2007 *Child labour, education and policy options,Division of policy and planning UNICEF* , (NY 10017)

[Damisa (2017)] Damisa,S., Musa,F., Sani, S., 2017, *Comparison of some link functions in binary response analysis*, *Biomedical Statistics and Informatics* . 2017; 2(4): 145-149

[Satriawan (2018)] Satriawan, E. & Ghifari,A,T ,2018, *How does parental income affect child labour supply? Evidence from the Indonesia family life survey.* ,*TNP2K, working paper 2- 2018* Australia government ,

[Laurent (2010)] Laurent, N. & Sebastien, D. 2010

Characteristics and determinant of child labour in Cameroon,CSAE Conference: Economic development in Africa $12^{st} - 23^{rd}$ March 2010, St Catherine's College , Oxyford. UK

[Lynn (1990)] Lynn, H. and Shlomo, S., 1990, Advanced calculus(new ed.),London England ,Jones and Bartlett International

[J2017] Jean, F. and Dhushyanth R. ,2017, Child labor across the developing world: Patterns and correlations, World Development Report

[UN (2015)] *The millennium development goals report 2015*, United Nations New York, 2015

[UN (2017)] *The sustainable development goals report 2017*, United Nations New York, 2017

[UNESCO (2005)] *Children out of School: Measuring exclusion from primary school*, UNESCO Institute for Statistics, Montreal , 2005.

[Moyi(2011)] Moyi, P. (2011) *Child labor and school attendance in Kenya.*,Educational Research and Reviews Vol. 6(1), p. 26-35

[Scott] cott, A. C. *Maximum Likelihood Estimation of Logistic Regression Models: Theory and Implementation*

[Long(1997)] Long, J. S. (1997) *Regression models for categorical and limited dependent variables:Advanced quantitative techniques in the social sciences (number 7)*, SAGE Publications Inc. London, U.K.

[Gujarati, D. (2007)] Gujarati, D. (2007) *Basic Econometrics*, Fourth Edition, Tata McGraw Hill education private limited, New Delhi.

[Nelder, J. A (2007)] Nelder, J. A. and Wedderburn, R. W. M. (1972) *Generalized Linear Models*, Journal of Royal Statistical Society, ‘series A, Vol. 135, P. 370 − 384.

[ILO(2013)] cott, A. C. *Global child labour trends 2008 to 2012* International Labour Office, International Programme on the Elimination of Child Labour (IPEC) – Geneva.

[SSCC (2019)] Social Science Computing Cooperative. *Supporting Statistical Analysis for Research*, The University of Wilconsin Madison , Madison.

[Kenward (1997)] Kenward and Rogers (1997).

APPENDIXES

Appendix I

Generalized linear model are of the form

$y_i = \beta_0 + \beta_{i1}x_1 + \beta_2 x_{i2}..., \beta_k x_{ik} + e_i$. i is the number of observation and $i = 1,2,3,...,n$,$y_i$is the response variable for observation i,j is the predictor variable and $j = 1,2,3,...,k$. y_i is a continours variable that has normal distribution. $Y_i \backsim N(\mu_i, \sigma^2)$
The expected value of response variable is

$$E[Y_i \mid x_1, x_2, x_3, ...,x_k] = \mu_i = \beta_0 + \beta_{i1}x_1 + \beta_2 x_{i2}..., \beta_k x_{ik}$$

The assumptions of general linear model are,

1. The Y_i are mutually independent normal random variables, with mean μ_i and constant variance σ^2. $Y_i \backsim N(\mu_i, \sigma^2)$.
2. Explanatory variables provide a set of linear predictors. $\eta_i = \beta_0 + \beta_{i1}x_1 + \beta_2 x_{i2}..., \beta_k x_{ik}$
3. The link between μ_i and η_i is that $\mu_i = \eta_i$. The mean of the dependent variable for any observation is the linear predictor formed from that observation's values on the explanatory variables.

This model may be unsatisfactory in a given practical situation in the following two situations:

1. When the distribution of Y_i is not normal.
2. When the mean of the dependent variable is a function of the linear predictor, rather than just the linear predictor itself.

To overcome pitfalls ,in 1972 Nelder and Wedderburn introduced the class of the generalized linear models (glm). The glm go beyond this in two major aspects:

1. Other than normal distribution, the response variables can have any other distribution within a class of distributions known as "exponential family of distributions".

2. we use a transformed function of the mean of predictor for the mean of the predictor.

$$g(\mu_i) = \beta_0 + \beta_{i1}x_1 + \beta_2 x_{i2} \ldots, \beta_k x_{ik}$$

for

$$\mu_i = \beta_0 + \beta_{i1}x_1 + \beta_2 x_{i2} \ldots, \beta_k x_{ik}$$

The exponential Family of Distributions

A random variable Y has a distribution within the exponential family if its probability density (or mass) function ($f(y)$) can be written in the canonical form $f(y, \theta, \phi) = e^{\{dfracy\theta \quad - \quad b(\theta)a(\phi) \quad + \quad c(y,\phi)\}}$.This function has three components that can be expressed as $f(y : \theta) = exp(a(y)b(\theta) \quad + \quad c(\theta) \quad + \quad d(y)$, where $a(y), b(\theta) and c(\theta)$ are function of response variable, unknown parameters and random variables respectively and $d(y) = ln1)$. $f(y, \theta, \phi)$ is a density function and therefore; $\int exp\{\frac{y\theta \quad - \quad b(\theta)}{a(\phi)} \quad + \quad c(y, \phi)\}dy \quad = \quad 1$.

Determining the first derivative of the probability function with respect to θ and simplifying we get;

$$\int \frac{[y \quad - \quad b^{'}(\theta)]}{a(\phi)} exp\{\frac{y\theta \quad - \quad b(\theta)}{a(\phi)} \quad + \quad c(y, \phi)\}dy \quad = \quad 0$$
$$\int [y \quad - \quad b^{'}(\theta)] exp\{\frac{y\theta \quad - \quad b(\theta)}{a(\phi)} \quad + \quad c(y, \phi)\}dy \quad = \quad 0$$
$$\int [yexp\{\frac{y\theta \quad - \quad b(\theta)}{a(\phi)} \quad + \quad c(y, \phi)\}dy \quad - \quad b^{'}(\theta) exp\{\frac{y\theta \quad - \quad b(\theta)}{a(\phi)} \quad + \quad c(y, \phi)\}dy \quad = \quad 0$$
$$\int [y\ quadexp\{\frac{y\theta \quad - \quad b(\theta)}{a(\phi)} \quad + \quad c(y, \phi)\}dy \quad = \quad b^{'}(\theta) \int \quad exp\{\frac{y\theta \quad - \quad b(\theta)}{a(\phi)} \quad + \quad c(y, \phi)\}dy$$
$$Compairing \quad the \quad two \quad sides \quad of \quad the \quad equation \quad we \quad get;$$
$$E(Y) = b^{'}(\theta) \tag{15}$$

Determining the second derivative of the probability function with respect to θ and simplifying we get;

$$\int \frac{-b^{''}(\theta)}{a(\phi)} exp\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\}dy + \int \frac{[y - b^{'}(\theta)]^2}{a^2(\phi)} exp\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\}dy = 0$$

$$\frac{b^{''}(\theta)}{a(\phi)} \int exp\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\}dy = \frac{[y - b^{'}(\theta)]^2}{a^2(\phi)} \int exp\{\frac{y\theta - b(\theta)}{a(\phi)} +$$

$$compairing \quad the \quad two \quad sides \quad we \quad get;$$

$$\frac{b^{''}(\theta)}{a(\phi)} = \frac{[y - b^{'}(\theta)]^2}{a^2(\phi)}, \quad since \quad b^{'} = E(Y), \quad [y - b^{'}(\theta)]^2 = var(Y) \quad hence;$$

$$\frac{var(Y)}{a^2(\phi)} = \frac{b^{''}(\theta)}{a(\phi)}, \quad \frac{var(Y)}{a(\phi)} = b^{''}(\theta),$$

$$\therefore$$

$$var(Y) = a(\phi) \cdot b^{''}(\theta) \tag{16}$$

Appendix II

Generalized likelihood ratio test statistic

The likelihood ratio test statistic is calculated from the MLE function described in eq(3.10) by

$$\Lambda = \frac{L_R}{L_S} = \frac{\Pi_{i=1}^{n} exp\{\frac{y\hat{\theta}_i - b(\hat{\theta}_i)}{a(\phi)} \quad + \quad c(y,\phi)\}}{\Pi_{i=1}^{n} exp\{\frac{y\widehat{\theta}_i - b(\widehat{\theta}_i)}{a(\phi)} \quad + \quad c(y,\phi)\}}.$$

where L_R is the MLE for the nested model , L_S is the MLE for the satulated model, $\hat{\theta}_i$ and $\widehat{\theta}_i$ are the estimate of i^{th} canonical parameters for the nested and saturated models respectively. $a(\phi) = \phi = 1$
simplifying these equation;

$$\begin{aligned} \Lambda &= exp \quad \Sigma_{i=1}^{n}\{[y\hat{\theta}_i - b(\hat{\theta}_i] - [y\widehat{\theta}_i - b(\widehat{\theta}_i]\} \\ &= exp \quad \Sigma_{i=1}^{n}\{[y(\hat{\theta}_i - \widehat{\theta}_i] - b(\hat{\theta}_i) + b(\widehat{\theta}_i]\} \end{aligned} \tag{17}$$

Since Deviance $D = -2Log\Lambda$ we take $-2nationallogarithm$ on both sides of the equation above. This produces

$$\begin{aligned} -2log\Lambda &= 2 \quad \Sigma_{i=1}^{n}\{[y(\widehat{\theta}_i - \hat{\theta}_i] - b(\widehat{\theta}_i) - b(\hat{\theta}_i]\} \\ D &= 2 \quad \Sigma_{i=1}^{n}\{[y(\widehat{\theta}_i - \hat{\theta}_i] - b(\widehat{\theta}_i) - b(\hat{\theta}_i]\} \end{aligned} \tag{18}$$

The Deviance of the model is indicated in Eq(3.9) where , $\theta_i = log(\frac{\pi_i}{1-\pi_i}) = \sigma_{j=0}^{p} x_{ij}\beta_j$, $b(\theta_i) = -nlog(1-\pi_i)$, and $\pi_i = \frac{e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}}{1+e^{\Sigma_{j=0}^{p} x_{ij}\beta_j}}$. $\therefore$ Deviance is the generalized likelihood ratio test statistic for comparing reduced model to the saturated model and nested model to the fitted model in mixed model .

Appendix III

R - Codes for the model

 # packages required

```
=require(ggplot2)
=require(GGally)
=require(reshape2)
=require(lme4)
=require(compiler)
=require(parallel)
=require(boot)
=require(lattice)
=library(nlme)
```

#Import data set in the csv file named X

```
ds<-read.csv("X.csv", header=TRUE, sep=",");
attach(ds)
#detach(ds)
str(ds)
names(ds)
```

#data visualisation

```
=ggpairs(pds[, c("TResidence", "hhsize", "Fhead", "HighestGradeEverAttended")])
```

#CrossTabulations

```
=library(gmodels)
=CrossTable(TResidence, Fhead)
=CrossTable(TResidence, hhsize)
=mytab = table(y, x)
=addmargins(mytab)
=install.packages("prettyR")
=library(prettyR)
=xtab(y   x, data=ds)
=boxplot(hhsize, horizontal=FALSE) # normally distributed
=boxplot(Age, horizontal=FALSE)# normally distributed
=boxplot(Average.monthly.cash.income.from.main.work, horizontal=FALSE) # normal, out-
liers on upper part
=boxplot(HighestGradeEverAttended, horizontal=FALSE)# skewed to right
```

=boxplot(HHNetMonthlyIncome, horizontal=FALSE)# normal
=boxplot(dsAge)
$= hist(ds$hhsize)

 # Descriptive statistics (with Pearsons Chisquare)
=chisq.test(ï..StatusOCLabourr, TResidence)# significant
=chisq.test(ï..StatusOCLabour, hhsize)# significant
=chisq.test(ï..StatusOCLabour, Fhead) #significant
=chisq.test(ï..StatusOCLabour, Sex) #significant at alpha = 10%
=chisq.test(ï..StatusOCLabour, Age) # significant
=chisq.test(ï..StatusOCLabour, PermanentDwelling) # not significant

 # model generation
#fit1 with all predictors (the saturated model)
=fit1 <- glmer(ï..StatusOCLabour   Age + Sex + hhsize + fhead+Tresidence
+ RelationHhh + HGradeAttended + Hours.spent.on.Household.chores + HHMonthlyIncome+
HHMonthlyExpenditure + (1|District) + (1|Division)
, data = ds,REML=FALSE,family=binomial(cloglog),control =
glmerControl(optimizer = "bobyqa"))
=fit1
= plot(fit1)
=summary(fit1)
=confint(fit1) # use p-value or CI to determine signficant of parameter
=anova(fit1)
= confint(fit1) # use p-value or CI to determine signficant of parameter
=anova(fit1)

#fit11 with one less fixed predictors (Age)(the nested model)
=fit11 <- glmer(ï..StatusOCLabour   Sex + hhsize + fhead+Tresidence +
RelationHhh + HGradeAttended + Hours.spent.on.Household.chores + HHMonthlyIncome+
HHMonthlyExpenditure + (1|District) + (1|Division)
, data = ds,REML=FALSE,family=binomial(cloglog),control =
glmerControl(optimizer = "bobyqa"))
= summary(fit11)
= logLik(fit1)
=AIC(fit1)

#testing overall significance of fixed effect model using LRT and anova for the two models.
= anova(fit11,fit1,test="Chisq")

#Testing significant of random effect
=fit1 # 2 random effect
#fit21 # one random effect
=fit21 <- glmer(ï..StatusOCLabour   Age + Sex + hhsize + fhead+Tresidence + RelationHhh + HGradeAttended + Hours.spent.on.Household.chores + HHMonthlyIncome+
HHMonthlyExpenditure + (1|District)
, data = ds,REML=FALSE,family=binomial(cloglog),control =
glmerControl(optimizer = "bobyqa"))
=summary(fit21)
=logLik(fit21)
LRT calculated using the loglik() function
=lrt = -2 * logLik(fit1) + 2 * logLik(fit21)
=pchisq(as.numeric(lrt), df=1, lower.tail=F)

futher model checks
=summary(fit1) # display results
=coef(fit1) #confint(fit1) # 95% CI for the coefficients
=exp(coef(fit1)) # exponentiated coefficients
=exp(confint(fit1)) # 95% CI for exponentiated coefficients
=predict(fit1, type="response") # predicted values
=residuals(fit1, type="deviance") # residuals
=plot(residuals(fit1, type="deviance"))

=detach(ds)