Statistical Analysis of Factors Associated With Early Marriage Among Women in Ethiopia Using Multilevel Logistic Regression Model

BY:

Lata Fekadu

A Thesis to be submitted to the School of Graduate Studies Jimma University Department of Statistics in the Partial Fulfillment of the Requirements for the Degree of Master of Science (MSc) in Biostatistics

February, 2020

Jimma, Ethiopia

Statistical Analysis of Factors Associated With Early Marriage among Women in Ethiopia Using Multilevel Logistic Regression Model

BY:

Lata Fekadu

Main Advisor: **Agatamudi Lakishmanarao (PhD)**

Co-Advisor: **Kibrealem Sisay (MSc)**

February, 2020

Jimma, Ethiopia

APPROVAL SHEET

JIMMA UNIVERSITY

SCHOOL OF GRADUATE STUDIES, DEPARTMENT OF STATISTICS

As thesis research advisors, we herby certify that we have read the thesis prepared by Lata Fekadu under our guidance, which is entitled **"Statistical Analysis of Associated Factors of Early Marriage Among Women in Ethiopia (Application of Multilevel logistic Regression Model)"** has been approved for submission to the Graduate Programs of Jimma University in partial fulfillment of the requirements for the award of the Degree of Master of Science in Statistics (Biostatistics) assembles with the regulations of the University and meets the accepted standards with respect to originality and quality.

Name of Advisors	**Signature**	**Date**
Agatamudi lakishmanarao (PhD)	____________	______________________
Kibrealem Sisay (MSc)	_____________	______________________

As the members of the board of examiners of MSc. thesis open defense examination. We certify that we have read and evaluated the thesis and examined the candidate. We recommend that the thesis be accepted as it fulfills the requirements for the degree of Master of Science in Biostatistics.

________________________________ __________________ __________________

Name of chairman **Signature** **Date**

_________________________________ ____________________________________

Name of Internal Examiner **Signature** **Date**

_______________________________ ____________________________________

Name of External Examiner **Signature** **Date**

STATEMENT OF AUTHORS

I declare that this thesis is a result of my genuine work and all sources of materials used, for writing it, have been duly acknowledged. I have submitted this thesis to Jimma University in the partial fulfillment for the degree of Master of Science in Biostatistics. The thesis can be deposited in the University library to be made available to borrowers for reference. I solemnly declare that I have not so far submitted this thesis to any other institution anywhere for that award of any academic degree, diploma or\ certificate.

Brief quotations from this thesis are allowed without requiring special permission provided that an accurate acknowledgement of the source is made. Requisites for extended quotations for the reproduction of the thesis in whole or in part may be granted by the head of the department of statistics when in his or her judgment the proposed use of the material is for a scholarly interest. In all other instances, however, permission must be obtained from the author.

Lata Fekadu

Date: ___________________

Signature: _______________

February, 2020

Jimma, Ethiopia

ACKNOWLEDGEMENT

First of all, I thank the almighty GOD for His love, mercy and protection upon me. I am indebted to many people, without their encouragement and share of ideas this research would be half. Thanks to the following blessed people.

I would like to extend my deepest thanks to my major advisor Agatamudi lakishmanarao (PhD), for his invaluable comments and suggestions that contributed to the successful realization of the thesis. I also gratefully acknowledge my co-advisor Kibrealem Sisay (MSc) for his continuous and passionate support by giving me constructive comments and suggestions which made me to concentrate on my thesis. My stay in the university was also a great pleasure and I am highly indebted to all of the staff members of department of statistics.

Also my special thanks go to Aboma Temesgen (Ass't Prof) who give me essential suggestions and future directions that really gives me a great motivation concerning the multilevel model. I am very much indebted to my friends and family members, especially to my Mother Kibitu Fayisa and my Father Fekadu Tolera who highly contributed to my life by giving me motivations as well as for their financial support starting from my childhood to today.

TABLE OF CONTENTS

TITLE **PAGE**

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AIC	Akaike Information Criteria
BIC	Bayesian Information Criteria
BDHS	Bangladesh Demographic and Health Survey
CSA	Central Statistics Agency
EAs	Enumeration Areas
EDHS	Ethiopian Demographic and Health Survey
ICC	Intra Class Correlation Coefficient
ICRW	International Center for Research on Women
IGLS	Iterative Generalized Least Square
IOM	International Organization for Migration
IPPF	International Planned Parenthood Federation
ICRWF	International Center for Research on Women Found
MDG	Millennium Development Goals
NCHTPE	National Committee on Harmful Traditional Practices in Ethiopia
PQL	Penalized Quasi Likelihood
MQL	Marginal Quasi Likelihood
STI	Sexual Transmitted Infection
UNDP	United Nation Development Program
UNFPA	United Nations Fund for Population Activities
UNICEF	United Nations Children’s Fund

ABSTRACT

Background: Early marriage is still widely practiced in many parts of the world mainly in developing countries which is particularly dominant in Sub-Saharan African country. Ethiopia has one of the highest rates of early marriage in the world as well as in Sub-Saharan Africa.

Objectives: This study is aimed to analyze associated factors of early marriage among women in Ethiopia with consideration of regional variability using multilevel modeling approach.

Methodology: The study is made based on the EDHS, 2016 data that has two-stages sampling hierarchical structure, collected for 9825 married women nested within eleven regions with age group 15-49 years. Descriptive statistics, single and multilevel logistic regression model analysis were used to identify determinants of early marriage and its variation across regional states.

Results: The results of the study showed that, of 9825 married women considered, 60.8% women were married at early age while 39.2% were married at an age of 18 years and above. The study also identified the major significant factors that affect early marriage among women. As a result, place of residence, religion of respondent, women education attainment, wealth index, husband education attainment, husband occupation status and total number of sibling were found statistically significant at 5% significance level. Furthermore, women who resides in rural area were (OR = 1.239) times more likely to be early married than those lived in urban. The reason behind is that urban women were more educated than those in rural implying that education was an important tools in delaying age at first marriage. The variance of the random component model related to the intercept term is statistically significant; implying the proportion of women got early marriage varies across regions have been accounted by random intercept terms only.

Conclusion: Among the three multilevel logistic models the random intercept model found to be the best fitting to the data. Thus, we conclude that those significant factors of women early marriage helps to implement more effective planning policies that target particular units at regional level. Particularly, education and wealth index have positive effect in reducing early marriage. As a result special attention needs to be paid to all regions in order to access education and improve the economic status for young women that help them to reduce early marriage.

Key Words: Multilevel model; Intra-class correlation coefficient; early marriage; logistic Regression model

1. INTRODUCTION

1.1 Background of the Study

Marriage is a moment of celebration and a milestone in adult life. In most societies, it is among the most significant life events for both men and women, indicating the emergence to adulthood and the beginning of new family building (Robert & Rebecca, 2010). Indeed, it is a universal social institution through which an adult male and female generally involves in marriage relationship and acquires new social status as a husband and wife. But, such an important social institution brings a numerous problem to a couple especially women, when it happens at an early age (Nasrin and Rahman, 2012).

Early marriage, also known as Child marriage, which is defined as "any marriage carried out below the age of 18 years, before the girl is physically, physiologically, and psychologically ready to shoulder the responsibilities of marriage and childbearing"(UNIFPA, 2006). This practice is now and for awhile understood as harmful practice on the health, psychological, physiological and socio-economic well-being of young girls as well as for the newborns (ICF International, 2011).

It is estimated that more than 12 million women worldwide first cohabited with a partner before the age of 18 without their will and consent, with the vast majority of them living in developing countries (UNICEF, 2017). Thus, it is argued that women early marriage is one of the most traditional practices in the globe that play a great role in lowering the status of women and children in particular (ICRW, 2010). This is predominantly practiced in the rural and poor communities of the developing countries where young girls are regarded as economic burden and quickly married off to alleviate household expenses (Kyari and Ayodele, 2014).

The practice of early marriage is most common in South Asia, Sub-Saharan Africa and Latin America and the Caribbean, where 48%, 42% and 29% of women aged 15-24 marry before the age of 18 respectively. It is very delicate among the developing countries such as Ethiopia as a result of the intact and deepening tradition, religion and economic motives which are the major reasons for the persistence of early marriage practice in the country (Kanta, 2013). The trouble was perpetuated by poverty, a lack of education and economic opportunities and social customs that limit the rights of women and girls (Mengistu, 2017). Thus, due to this inhuman and discriminatory practice, women and girls are banned from access to education, health care services, employment and other opportunities and resources. It is also one of the global problems that undermine the personal, nations and countries development and the rights of women very seriously (Mengistu, 2017). Early marriage also threatens the achievement of MDGs such as eradicating extreme poverty and hunger, achieving universal primarily education, promoting gender equality and empowering women, reducing child mortality, improving maternal health and combating HIV/AIDs, malaria and other diseases (UN, 2007).

According to UNICEF, Ethiopia has the 15th highest prevalence rate of child marriage in the world. Similarly in Ethiopia, 40% of girls are married off before 18 years old. 14% are married before they turn 15. Ethiopia is the 16th highest nation in the world for child marriage (UNICEF, 2017).

Also according to EDHS (2016) report, the median age at first marriage among women age 25-49 has increased slightly since 2011 to 2016 from 16.5 years to 17.1 years. During the same period, the percentage of women marrying before age 18 has declined from 63% to 58%. However this proportion is not the same throughout the country. The rates in Amhara and Tigray region are much higher than the national average (82% in Amhara, 79% in Tigray, 64% in Benshangul, 64% in Gambella and 46% in Afar region (NCTPE, 2003).

Multilevel regression models are increasingly applied in many areas of social and biomedical science data sets containing identifiable units or clusters of observations (Goldstein, 2003). Social research usually involves problem that investigate the relationship between individual and society. The general concept is that individual interacts with the social contexts to which they belong and the properties of those groups are in turn influenced by the individuals who make up that group. Generally the individuals and the social groups are conceptualized as a hierarchical system of individuals and groups. Naturally such systems can be observed at different hierarchical levels and variables may be defined at each level. This leads to research in to the interaction between variable characterizing individuals and groups respectively (Hox J., 2002).

A potential drawback to multilevel modeling is the additional complexity of coefficients varying by group. It does create new difficulties in understanding and summarizing the model (Gelman, 2006). Multilevel modeling methodology is applied to the hierarchically structured data in which the units at one level are clustered with the units of the next higher level. Multilevel model allows the simultaneous examination of the effects of group level and individual level variation dependence of observations within and between groups (Snijders and Bosker, 1999).

The data set used in this study was EDHS, 2016 which was based on two-stage stratified cluster sampling in which the individual observations are correlated. Multilevel analysis is a methodology for the analysis of such data with complex patterns of variability, with a focus on nested sources of variability. Here the units at lower level are married women who are nested within units at higher level (regions). Therefore the prevailing consequence of women early marriage and its variation due to nature of the data calls for intervention in view of identifying the determinants of early marriage and quantifying its variation among women in Ethiopia using multilevel model. This is what initiated the investigator to conduct the study related with women early marriage.

This thesis is organized in five sections. The statement of the problem, objective of the study and its significances are presented next in this section. Section 2 describes some literature review related early marriage among women in Ethiopia and multilevel modeling approach. In Section 3, the data and the detail methods of data analyses employed are explained. Then, basic results of the study and discussions of the results are presented under section 4. Finally, some concluding remarks and recommendations are provided in Section 5.

1.2 Statement of Problem

Early Marriage is one of the global alarming problems that undermine the personal development and the rights of women very seriously (Amin S., 2008). Also it has a direct oppose on realizing of MDGs (Myers & Harvey, 2011). In Ethiopia, two in every five girls are married before their 18th birthday and nearly one in five girls marries before the age of 15 (UNICEF, 2017).

In Ethiopia child marriage has devastating implications for the girls, family and the community as a whole. This is linked to reinforcing cycles of poverty, increased rates of maternal and infant mortality, gender inequality and low education rates for girls. Irrespective of the efforts of the government, the society and international community; the problem was still persistent throughout the country (Mengistu, 2017). In 2017 in Ethiopia, 40% of girls are married off before 18 years old which was still high (UNICEF, 2017).

Eshetu *et al.,* (2012) conducted the study to identify determinants of age at first marriage in Addis Ababa, Ethiopia. He used logistic regression model that limit to identify the variability of women early marriage among region due to the nature of data. Similar methodology was applied in Sinana District, Northwest Ethiopia to identify determinants of early marriage among women. The result showed that the odds of early marriage practice were 12.2 times higher among rural residents compared to urbanites and this cannot represent the overall of multi regional setting of Ethiopia (Sileshi *et al.,* 2015). Similarly chi-square and binary logistic was applied to investigate the spatiotemporal variation of early marriage in Babile woreda of East Hararge. Accordingly, current age, educational status, parental control, cultural as well as religious influences and ethnicity were the main factors significantly associated with early marriage in the study area (Mohammed, 2018).

Most early marriage study in Ethiopia has been small-scale research, focusing on some part of communities, usually small-sized rural or urban communities. Their geographic scope limits the applicability of their result on a large scale, particularly considering the complex multi-regional and multi-ethnic setting of Ethiopia. In addition, some researcher used logistic regression model and Cox proportional hazard model to estimate the effect of covariates on women early marriage which restricts consideration of regional variability of women early marriage when data are clustered type. However the researcher wants to use multilevel logistic regression model that permit analyzing the loss of independence observations turn out from clustering individual married women in to higher level (region). Likewise, it allows researcher to make valid inferences when examining the effect of both individual characteristics and cluster characteristics on the outcome of the response.

In doing so the aim of this study is to assess the within-region and between-region level variation of early marriage, to identify the true effect determinant factors of early marriage by applying multilevel model which also helps to implement more effective planning policies that target particular units at two levels of the hierarchy. The research questions are:

- What are the major factors that affect early marriage among women in Ethiopia?
- Does early marriage among women vary within and across regional states of Ethiopia?
- Are socio-economic and demographic factors responsible for variation of women early marriage among regions of Ethiopia?

1.3 Objectives of the Study

1.3.1 General Objectives

The main objective of this study is to analyze associated factors of early marriage among women in Ethiopia with consideration of regional variability using multilevel modeling approach.

1.3.2 Specific Objectives

- To identify the socio-economic and demographic determinant factors of early marriage among women in Ethiopia
- To assess the extent of variation in women early marriage between and within region of Ethiopia
- Examining the existence of factors that are responsible for variation of women early marriage between regions of Ethiopia.

1.4 Significance of the Study

The results of this study may help the organization as well as individuals who work in this area to get a clue on to what extent of women early marriage variation were between and within regions of Ethiopia which helps to the government in setting policies and strategies. It will also serve as a reference for giving intervention accordingly to conduct further researches and to make strategic recommendations. Additionally the study uses the researchers for understanding the analysis of hierarchical structure of clustered data. This helps to model and give emphasis on the factors that have strong association with women early marriage, so that policy makers act in accordingly. It may also be an input for the government to create awareness for the community by identifying the potential risk factors of women early marriage. The result of this study will also be expected to help those policy makers on harmful traditional practices. The researcher believe that working this study will boost up his experience to conduct skill full studies in his future career and take active part in various government and NGO activities to solve similar socio economic problems of the society.

1.5 Limitation of the Study

It is worthwhile to mention some of the constraints that the researcher faced while undertaking this study, the first constraint was the study focused on identifying factors that are expected to influence women early marriage in Ethiopia. However, due to lack of data the study could not incorporate some of the most influential factors such as cultural influence of the society and parental consciousness which were captured by qualitative type of data. Since prevention is better than statistical cures the problem of missing values and non response rate for some variables were the other limitations of this study. Similarly the data used in this study was the 2016 EDHS. Thus, the results may not necessarily reflect the current situation of women early marriage in Ethiopia.

2. LITRATURE REVIEW

2.1 General Review on the Women early Marriage

According to World Health Organization, health refers to a complete state of wellbeing, mental, physical and social and not merely the absence of diseases (WHO, 2006). This implies that certain social/cultural aspects affect the health of an individual. Though, the persistence of the early marriage is a way affecting the health of an individual or the community (UNICEF, 2014).

Each day, more than 41,000 girls worldwide are married while still children, often before they may be physically and emotionally ready to become wives and mothers. This endangers the life trajectories of girls in numerous ways. Such as greater risk of experiencing a range of poor health, having children at younger ages, having more children over their lifetime, dropping out of school, earning less over their lifetimes and living in poverty than their peers who marry at later ages (UNICEF, 2014).

Early marriage may also be more likely to experience intimate partner violence, have restricted physical mobility, and limited decision making ability. Most fundamentally, these girls may be disempowered in ways that deprive them of their basic rights to health, equality, non-discrimination, and to live free from violence and exploitation, which continue to affect them into adulthood. These dynamics affect not only the girls themselves, but their children, households, communities and societies, limiting their ability to reach their full social and economic potential (Wodon *et al.*, 2017).

Similarly the Convention on the Elimination of all forms of discrimination against Women as the study period estimates show that approximately 82 million girls in the world between 10–17 years will be married before they reach 18 years of the 331 million girls aged 10–19 in developing countries 163 million will be married before they are 20 (Fitch *et al.*, 2011). In India almost half (44.5%) of women aged 20-24 years got married before they reach 18 years where they grow up with the normative expectation of marriage within a socially determined social frame (Sinha, 2009).

Child marriage is continues to be highly prevalent in Africa, where almost 40% of girls marry before reaching 18 years (UNICEF, 2014). This prevalence also substantially varies from country to country. Accordingly; the highest is in Niger with 75% and followed by Chad and Central African Republic with 68%, Guinea with 63%, Mozambique with 56%, Mali with 55%, South Sudan 52%, Somalia 45% and Eritrea 41% which is highly prevalent in sub-Saharan Africa (ICRW, 2015).

Ethiopia has one of the highest rates of early marriage in Sub-Saharan Africa. The study conducted by Gordon (2012) found that despite the large differences in the proportion of females marrying before the ages of 15 and 18, on the far side of Addis Ababa, prevalence of early marriage is high across the country. The difference in prevalence across regional states ranges from 12% in Addis Ababa to 58% in Benishangul Gumuz; where as 41% Oromiya (Federal Democratic Republic of Ethiopia, 2012). Similarly study by the National Committee on Harmful Traditional Practices of Ethiopia (NCTPE, 2003) found that the estimated proportion of women married before the age of 15 is 57%. This practice is occurred in more extreme in the northern part of Ethiopia where girls are married as young as eight or nine years of age. In some instances, they are even pledged at birth.

2.2 Associated Factors of Women Early Marriage (Empirical Review)

Region of Residence

The women early marriage varies significantly by region of residence. Region of residence identifies the geographic regions in which the respondents were interviewed, which were classified in to nine regional states and two administrative cities. Studies in Ethiopia reveal that Ethiopia has one of the highest rates of female early marriage in the world, with one in two girls marrying before her 18th birthday and one in five girls marrying before the age of 15. However, prevalence rates vary greatly by region and are often higher than national figures, such as in the Amhara region in northern Ethiopia, where almost 50 percent of girls are married by age 15 (Assefa *et al.*, 2005)

Place of Residence

Based on the result of qualitative data which analyzed through thematic analysis state that one of the biggest arguments as to why child marriage still occurs in Ethiopia is because of its presence in history and traditional practices often seen in rural Ethiopia. It is important to understand that child marriage has been around since the beginning of Ethiopian civilization it has become normalized in rural society, even encouraged (Shiferaw.*et al.,* 2013).

The descriptive results of survey conducted in Ethiopia shows that among women married before age 15, 82% resided in rural areas. This is because of rural community slightly lower age cut-offs for defining early marriage for girls than urban. 66% of mothers knew that the legal minimum age at first marriage was 18. The most commonly recognized consequences of early marriage by care takers were increased poverty (54%), more obstructed labor (47%), high obstetric fistula (30%), higher maternal mortality (22 %), and less education for girls (21%) (Gage AJ., 2007).

Annabel Erulkar (2013) conducted the study to examine factors of early marriage among seven region of Ethiopian. He used Cross tabulation and logistic regression model and found that among women married before age 15, 82% resided in rural areas of Ethiopia and 79% had never been to school. Study conducted in Nigeria used Chi-square and Cox proportional hazard models to determine the survival time of age at first marriage among women of reproductive age in Nigeria using NDHS-2008. The result showed that the mean age at first marriage was 17.8 and place of residence was significantly associated with age at first marriage. I.e. women who reside in rural area (H.R=1.15) married earlier than their counterpart in urban area (Adebowale *et al*., 2012).

Using Shared Frailty Models Bedasa *et al.,* (2015) revealed that women in rural areas tend to have institutional and normative structures such as the kinship and extended family that promote early marriage and childbearing, but women in urban areas need to develop skills, gain resources, and achieve maturity to manage an independent HH that shows women who lived in urban areas are more survived on age at first marriage than women who lived in rural areas. Study conducted using logistic regression in Bangladesh showed that the vulnerability of children in rural to child marriage is still higher, compared to children reside in urban area (71% in rural areas as opposed to 54% in urban) (Sarker *et al.,* 2010).

Religion of Respondent

Religious institutions are key enforcers of social norms and have considerable potential to bring about changes in gendered social norms. On his research Sarker (2010) got another finding that Bangladesh is one of the largest Muslim countries in the world, where early marriages are widely practiced. Accordingly, 92% of Muslim women in the country married earlier compared with 84% of Christianity and 85% of Buddhism. Similarly, using binary logistic regression model study in Serbia by Hotchkiss *et al.,* (2016) revealed that, early marriage is high among non-Orthodox women than Orthodox women (55% versus 47%). In addition to this, pointing out the awareness of different religious followers on the early marriage, the survey conducted by Boyden *et al.,* (2013) in Ethiopia on the early marriage and associated factors revealed that the highest awareness was among protestant Christians (80%), and there was slightly higher awareness among orthodox Christians (74%) than among Muslims (70%).

Religion is often blamed for the prevalence of early marriage. The study conducted in Durban, South Africa using logistic regression model tries to know the cause and prevalence of child marriage. The researcher found that among 12 Sub-Saharan African countries certain religious affiliations were positively associated with child marriage where its prevalence was higher among women who practiced Islam, traditional religions or no religion than among women who were Christians (OR = 1.2 to 1.3) (Belinda, 2015).

The results of qualitative research in Nepal, Ethiopia and India, indicate that parents may marry off their daughter because they fear her being sexually active outside of marriage; perceive their daughter's value to be greater doing housework than studying for a job that does not exist; or to avoid paying the higher dowry that often comes with marrying off their daughter when she is older (Nanda, 2015; ICRW, 2017).

Women and Husband Education Level

Education is the power to challenge discriminatory social norms in societies where girls are not ascribed the same value as boys (UN, 2008). The study conducted in India revealed that, about 25% of those who cannot read and write females are married before 15 years of age, while 40% are married between 15 and 17 years, as opposed to only 1% and 5%, respectively, of women with post-secondary education (Sofia and Khalid, 2015). Similarly, girls with higher levels of schooling in Africa are less likely to marry as children. In Mozambique, for instance, some 60% of girls with no education are married by 18, compared to 10% of girls with secondary schooling and less than one % of girls with higher education (ICRW, 2015).

In Ethiopia, according to descriptive report of CSA (2012), age at first marriage greatly increases with education; women with more than secondary education get married almost eight years later than those with no education. Accordingly, evidences from (Alemayehu, 2014) show that women education has a substantial impact on early marriage. Thus, the results of logistic regression analysis revealed that in Ethiopia, women 20 to 29 years old with secondary or higher education marry at an average age 3.2 years higher than that of women of the same age with no education

Wealth Index

In almost all developing countries, child marriage is more common among the poorest people than the wealthiest (ICRW, 2015). Accordingly, Hotchkiss *et al.,* (2016) apply multiple logistic regression models in order to identify the risk factors of child marriage among girls in Serbia. The investigation shows that about 24.3% of females living in the poorest quintile of HHs were married by age 15, compared to 12.4% of those in the middle wealth group and 3% of those in the richest wealth groups. The study conducted in Senegal and Côte d'Ivoire shows that girls in the poorest 20% of HHs are more than four and three times as likely to be married as girls in the richest HHs (ICRW, 2010).

In Ethiopia context, female early marriage is seen as a way to improve the economic status of the family. The community-based cross sectional study conducted by Sileshi *et al.,* (2015) in Sinan Woreda of North West Ethiopia. Accordingly, logistic regression analysis resulted that families with monthly income of ranging 451-650 ETB were 2.5 times more likely to practice early marriage compared to those having monthly income of more than 800 ETB.

Other Factors

Empirically, many studies have shown that age at first marriage is influenced by a number of socio-economic and demographic factors. For instance Peninah *et al.,* (2011) investigated determinants of early marriage among women in western Uganda using Cox's proportional hazard model and the result showed that the educational attainment of women, religion, district of residence and husband occupation were the determinants of age at first marriage. The risk of women early marriage was 18% and 34% lower for the women with primary education and at least secondary education as compared with non educated women. The significance levels for all educational categories were significant and thus risk of getting early marriage reduced as the level of education increased. These results provide empirical evidence that a woman's educational attainment was an important determinant of a women's age at first marriage in Western Uganda.

Joseph *et al.,* (2012) try to examine the effect of socio-economic factors on prevalence of child marriage and its determinants among young women in Indonesia by using logistic regression. The result shows that there was a negative correlation of early marriage with higher income of households, exposure to the media through the internet, education of household head, and number of children in a family.

The study by Choe *et al.,* (2005) in Nepal used a proportional hazards model in order to examine the effect of covariates on early marriage. The result of the studies revealed that age at first marriage was varied by the ecological zones of the Hills, Mountains and Terai regions i.e., region of residence is significantly determine the marriage stage. In addition to region of residence, they also found that education was played an important role. The studies have found that children of parents with higher education were less likely to get married at an early age.

Zahangir and Kamal (2011) worked on several attributes linked with child marriage of females' in Bangladesh by using binary and multiple logistic regression models. They revealed that early marriages were more frequent among the women who are rural childhood, born in Muslim community, live in rural area, no/less educated, marry with no/less educated husbands, have no access to mass media, and have a lower economic status. It is argued that higher educational attainment was the main force underlying the delay in a first marriage among females.

Mohammed (2018) used qualitative and quantitative data to investigate the spatiotemporal variation of early marriage in Babile Woreda, using chi-square association for analysis. The results showed that family size, place of residence, religion of a woman, educational level of women, respondents work status, household's wealth index, exposure to mass media and occupation status were found to be significantly associated with women early marriage.

2.3 Multilevel Modeling Review

Langford (1993) and Goldstein (2003) describe that the multilevel is extension of generalized linear models. Multilevel analysis is a suitable approach to take into account the social contexts as well as the individual respondents or subjects. Thus it allows the simultaneous examination of the effects of group level (cluster) and individual level variables on individual level outcomes while accounting for the non-independence of observations within groups. Also this analysis allows the examination of both between group and within group variability as well as how group level and individual level variables are related to the response variable (Khan and Shaw, 2011).

Why we need multilevel modeling? A multilevel problem concerns a population with a hierarchical structure. A sample from such a population can be described as a multistage sample: First, the sample of units was taken from the higher level units (e.g., schools), and next the sample of sub-units was taken from the available units (e.g., the sample of pupils from the schools). In such samples, the individual observations are in general not completely independent. For instance, pupils in the same school tend to be similar to each other, because of selection processes (for instance, some schools may attract pupils from higher social economic status levels, while others attract lower social economic status pupils) and because of the common history the pupils share by going to the same school. As a result, the average correlation (expressed in the so-called intra-class correlation) between variables measured on pupils from the same school will be higher than the average correlation between variables measured on pupils from different schools (Hox J., 2010).

Multilevel models are extensions of generalized linear model in which data are structured in groups and coefficients can vary by group. The advantage of multilevel model over the single logistic regression is that, in classical regression ignoring group indicators can be misleading the group-level variation. Multilevel modeling allows the estimation of group averages and group-level effects, compromise between the overly noisy within-group estimate and the over simplified regression estimate that ignores group indicators (Gelman *et al.,* 2006).

With grouped data, a regression that includes indicators for groups is called a varying-intercept model because it can be interpreted as a model with a different intercept within each group. This is the case where random intercept model is considered. Which mean that the group differ with respect to the average value of individual level, but there is no different relation between indicators of response among groups (regional states) (Gelman *et al.,* 2006).

In EDHS, the structure of data in the population is hierarchical, and a sample from that population can be viewed as a two-stage sample. Because of different cases like, cost, time and efficiency considerations, stratified multistage samples are the norm for sociological and demographic surveys. Cluster sampling system often introduces multilevel dependency or correlation among the observations (married women) that can have implications for model parameter estimates. The problem of dependencies between individual observations (married women) also occurs in survey research, where the sample is not taken randomly but cluster sampling from geographical areas is used instead. In this case, the use of single-level statistical models is not reasonable. Hence, in order to draw appropriate inferences and conclusions from multistage stratified clustered survey data, we may require multilevel modeling.

The multilevel modeling strategy accommodates the hierarchical nature of the DHS data and corrects the estimated standard errors to allow for clustering of observations within units (Goldstein, 2003). The fact that the regional states of Ethiopia had a variety of environmental factors like place of residence, women and partner's education level, economic status and access to any media that encourage the reduction of women early marriage at their region as well as national level (Sileshi *et al.,* 2015, Mohammed, 2018). Indeed, not only regional-level differentials but also there are the individual-level factors attributed for women early marriage as well. This differential among individual and regional level indicated the facts that, the rate of women early marriage in Ethiopia has different structure (Bedasa *et al.,* 2015). But, so many studies in single level (eliminating those variations across regional states) regarding women early marriage in Ethiopia and worldwide that invites errors. In fact, there is clear heterogeneity among the individual and regional-level. In such situation multilevel logistic regression is the natural choice for modeling.

3. DATA AND METHEDOLOGY

3.1 Source of Data

This study used Demographic and Health Survey data conducted in Ethiopia in 2016, which was the fourth comprehensive survey conducted as part of the worldwide Demographic and Health Surveys project. The Ethiopia Demographic and Health Survey were implemented by the Central Statistical Agency (CSA) and partner organization under the auspices of the Ministry of Health from January 18, 2016, to June 27, 2016. The data provide in-depth information on marriage, fertility, family planning, infant, child, adult and maternal mortality, maternal and child health, gender, nutrition, malaria, knowledge of HIV/AIDS and other sexually transmitted diseases.

3.1.1 Sample Design

The sample for the EDHS, 2016 was designed to provide population and health indicators at the national (urban and rural) and regional levels. The 2007 Population and Housing Census, conducted by the CSA, provided the sampling frame from which the EDHS 2016 sample was drawn. Administratively, Ethiopia is divided into nine federal regions and two administrative cities. The sample for the EDHS, 2016 was designed to provide estimates of key indicators for the country as a whole. The EDHS, 2016 sample was selected using a stratified and in two-stage sampling. Each region was stratified into urban and rural areas, giving 21 sampling strata. Samples of EAs were selected independently in each stratum in two stages.

In the first stage, a total of 645 EAs including 202 EAs in urban areas and 443 EAs in rural areas were selected with probability proportional to the EA size based on the 2007 PHC. A household listing operation was carried out in all the selected EAs from September to December 2015. In the second stage of selection, a fixed number of 28 households per cluster were selected with an equal probability systematic selection. This two-stage sampling EDHS, 2016 data set is of hierarchical structure. The hierarchy for this study follows individuals/married women as level-1, and regions as level-2 (i.e. individuals women are nested in to regions). The study population in this research was all women within the reproductive age group (15-49) years that were ever married or ever-lived with a man as a wife, living in Ethiopia during the study of the EDHS, 2016.

Target Population

In the DHS all women were asked a series of questions regarding their marital status and whether they had ever married or lived with a man or not. All those who reported that they were ever married or ever-lived with a man, were asked to indicate how old they were at the time when they started, for the first time ever, living with a man as a wife, irrespective of the legality or otherwise of their union. The response to this question constitutes the woman's age at first marriage. All the women who indicated that they had never been in a union or lived with a man were considered single and as a result they were not asked the question about the age at first marriage. This is the standard way in which age at first marriage is being measured in the worldwide DHS program (Ikamari, 2005). Missing value is common in EDHS data. For this study there are few variables which have some missing values. Thus, after clearing those missing values, the analysis of the study was presented on 9825 married women aged 15-49 among 15683 all interviewed women on EDHS, 2016.

3.2 Variables of the Study

3.2.1 Response Variable

One question from the EDHS used to examine the dependent variable, the age of women at first marriage which is either less than 18 years (early marriage) or 18 years and above (legal marriage). The response variables were dichotomous, coded as 0 if age of first marriage were 18 years and above and coded as 1 if age of first marriage were less than 18 years. Therefore the response variable for the i^{th} married women is represented by a random variable Y_i with two possible values 1 or 0 and coded as follows:

$$Y_i = \begin{cases} 1, \text{ if age at first marriage is less than 18 years} \\ 0, \text{ if age at first marriage is 18 and above years.} \end{cases}$$

3.2.2 Independent Variables

The primary choice of explanatory variables for this study was based on literature reviews and theoretical justification of source of data on factors influencing women early marriage at the global level and in the country. Therefore, those variables, which are reviewed from literature as determinant factors of women early marriage, were displayed on Table 3.1.

Table 3.1 The independent variables and values of the category

Variables	Values of category		
1. Region of residence	(1) Tigray (ref.) (2) Afar (3) Amahara (4) Oromia (5) Somali (6) Ben-gumes (7) SNNP (8) Gambella (9) Harari (10) Addis Ababa (11) Dire Dawa		
	Values of category	Variables	Values of category
2. Place of residence	(1) Urban (ref.) (2) Rural	6. Number of sibling	(1) 5 & less (ref.) (2) above 5
3. Religion of a woman	(1) Orthodox (ref.) (2) Muslim (3) Protestant (4) Others	7. Husband's occupation	(1)Agriculture (ref.) (2)Professional (3) Business (4) Labourers (5) Others
4. Women's education	(0) None (ref.) (1) Primary (2) Secondary (3) Higher	8. Husband's education	(0) None (ref.) (1) Primary (2) Secondary (3) Higher
5. Wealth index	(1) Richest (ref.) (2) Richer (3) Middle (4) Poorer (5) Poorest	9. Respondents work	(0) Working (ref.) (1) Not working
		10. Media Exposure	(1) Yes (ref.) (2) No

3.3 Method of Data Analysis

This study was exploring EDHS data conducted in 2016 related to women early marriage that have binary response variables. A range of techniques has been developed for analyzing data with categorical response variables. For this study, the extensions of generalized linear model like multilevel logistic regression model applied with the help of R statistical software.

3.3.1 Introduction to Generalized Linear Model

Generalized linear models (GLM) extend ordinary regression models to encompass non-normal response distributions and modeling functions of the mean. Three components specify a generalized linear model: A random component identifies the response variable Y and its probability distribution; a systematic component specifies explanatory variables used in linear predictor function and a link function specify function of E(Y) that model equates to the systematic component. Therefore the GLM generalizes linear regression by allowing the linear model to be related to the response variable via a link function (Agresti, 2002).

Binomial Distribution

Often, categorical data results from n independent and identical trials with two possible outcomes for each, referred to as "success" and "failure". Let y_i denote the number of early married women out of n number of married women. Under the assumption of n independent and identical trials, Y_i is the random variable having the binomial distribution with index n and parameter π_i, where π_i the probability of success (women early married) and denoted by:

$$Y_i \sim B(n, \pi_i).$$

The probability density function of binomial distribution for the outcome y_i from random variable Y_i is given by:

$$P\{ Y_i = y_i\} = \binom{n}{y_i} \pi_i^{y_i}(1-\pi_i)^{n-y_i} \text{ for } i = 0,1,2,\dots,n \tag{3.1}$$

An important property of the GLM is the functional relation between mean and variance. That means any factor that affects the probability of success will alter both mean and variance of the observations simultaneously. This suggests that a linear model that allows the predictors to affect the mean but assumes that the variance is constant will not be adequate for the analysis of binary data (G. Rodr´ıguez, 2007).

Odds and Odds Ratio

Odds are the ratio of probability of an event will occur divided by the probability of it will not occur. In this study, the event is that the women i married before 18 years old $y_i = 1$ and given by:

$$\text{Odds (Early Married)} = \frac{P(EM)}{P(\text{not EM})} = \frac{P(EM)}{1-P(EM)} \tag{3.2}$$

Where, EM is early marriage and $\mathrm{P(EM)}$ is the probability of women earl marriage.

Odds always have values greater than zero and in this study if odds value is larger than one the probability of women early marriage will occur more likely than that of legal marriage. Odds ratio, as the name indicates, is the ratio of two odds and given as follows:

$$\mathrm{OR} = \left(\frac{\mathrm{P_1(EM)}}{1-\mathrm{P_1(EM)}}\right) / \left(\frac{\mathrm{P_2(EM)}}{1-\mathrm{P_2(EM)}}\right) \tag{3.3}$$

Here, $\mathrm{p_1}$ and $\mathrm{p_2}$ refers to the probability of women early marriage in group one and group two respectively. If the odds ratio value is greater than one indicates that the odds of the outcome in group one is larger than in group two. Thus, married women in group one is more likely to have early married than married women in group two. In binary logistic regression analysis, odds ratio is the exponent of the estimated coefficient $\hat{\beta}$, $\exp(\hat{\beta})$.

3.3.2 Logistic Regression Model

Regression methods have an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables. It is often the case that the outcome variable is categorical, taking on two or more values. When the outcome variable is binary or dichotomous many distribution functions have been proposed for use. Logistic regression model can be used mainly for two reasons. The first is from a mathematical point of view, it is an extremely flexible and easily used function, and the second it leads itself to meaningful interpretation (Hosmer and Lemeshow, 2011). The assumptions required for statistical tests in logistic regression are far less restrictive than those for ordinary least squares regression.

Let Y be a dichotomous outcome random vector with categories 1 (women early married) and 0 (women legal married). Let X be an n x (k+1) matrix that contains the collection of k-predictor variables of Y. Then, the conditional probability that the $\mathrm{i^{th}}$ married women experiences early marriage given married women characteristics $\mathrm{X_i}$ is given by:

$$\pi_\mathrm{i} = \mathrm{P(y = 1/X_i)} \tag{3.4}$$

In logistic regression analysis, it assumed that the explanatory variables affect the response through a suitable transformation of the probability of the success. This transformation is a suitable link function of π_i , and is called the logit-link, which is defined as:

$$\text{Logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} \quad (3.5)$$

Where $\beta = (\beta_0, \beta_1, \beta_2, \ldots, \beta_k)'$ are the model parameters and $X_i = (X_{0i}, X_{1i}, X_{2i}, \ldots, X_{ki})$ with $X_{0i} = 1$ and $i = 1, 2, \ldots, n$ are factors associated with women early marriage.

Therefore the probability of success expressed as:

$$\pi_i = \frac{e^{X'_i\beta}}{1+e^{X'_i\beta}} = \frac{1}{1+e^{-X'_i\beta}} \quad (3.6)$$

With further rearrangement, we obtain the odds of success

$$P(y = 1/X_i) = \frac{\pi_i}{1-\pi_i} = e^{X'_i\beta} \quad (3.7)$$

3.3.2.1 Maximum Likelihood (ML) Estimation of the Parameters

The logistic regression model just described is a generalized linear model with binomial errors and logit link. Thus the aim of analyzing logistic regression is to estimate the k+1 unknown parameters β'^s from equation (3.5). The most commonly used methods of estimating the parameters of logistic regression model are the method of maximum likelihood (ML) estimation. The method of maximum likelihood estimation yields to estimate values for the unknown parameters which maximize the probability of obtaining the observed set of data.

Suppose we have a sample of n independent observations (y_i, x_i), i=1, 2, …, n. Where y_i denotes the values of a dichotomous outcome variable and X_i is the value of the explanatory variables for the i^{th} married women.

Assume that $Y_i \sim \text{Bernoulli}\ (1, \pi_i)$ and the probability mass function of Y_i are given by:

$$f_i(Y_i) = \pi_i^{y_i}(1 - \pi_i)^{1-y_i} \quad (3.8)$$

We define the likelihood function as the joint probability density function in Equation 3.8 expresses the values of y as a function of known, fixed value of β. That is:

$$L(\beta|Y) = \prod_{i=1}^{n}(\pi_i^{y_i}(1 - \pi_i)^{1-y_i})$$

$$L(\beta|Y) = \prod_{i=1}^{n} \left[\frac{\pi_i}{1-\pi_i}\right]^{y_i} [1-\pi_i]$$

The maximum likelihood estimates of the parameters are obtained by maximizing the log-likelihood function which is given by taking the natural logarithm of both sides yields the following expression for log likelihood function:

$$\mathrm{LogL}(\beta|Y) = \sum_{i=1}^{n} \left[y_i \ln\left[\frac{\exp(x_i{}'\beta)}{1+\exp(x_i{}'\beta)}\right] + (1-y_i)\ln\left[\frac{1}{1+\exp(x_i{}'\beta)}\right]\right] \quad (3.9)$$

The maximum likelihood estimates of the parameters are found by the derivation of the log-likelihood function with respect to each $\beta's$ and set each equation to zero which is given as:

$$\frac{\partial \ln L(\beta|Y)}{\partial \beta_j} = \frac{\partial\left(\sum_{i=1}^{n}\left[y_i \ln\left[\frac{\exp(x_i{}'\beta)}{1+\exp(x_i{}'\beta)}\right] + (1-y_i)\ln\left[\frac{1}{1+\exp(x_i{}'\beta)}\right]\right]\right)}{\partial \beta_j} = 0$$

$$\text{for } j = 1, 2, \dots, k$$

3.3.2.2 Model Building and Variable Selection

The number of variables to be included in the model should be the minimum possible that is parsimonious and deliver optimum information. In this study the variable selection process begins with univariable analysis of each independent variable with response variable separately. Tests to determine whether a systematic relation or association between each predictor variable with the response variable exists are made before the final model was selected. Upon the completion of the univariable analysis, we select variables for the multiple logistic regression analysis. Any variable whose univariable test has a p-value ≤ 0.25 is a candidate for multiple logistic regression model along with all variables of known statistically importance. (Keith and David, 2007)

Another approach to variable selection is to use stepwise selection procedure. In this method, variables are selected for either inclusion or exclusion from the logistic regression model in a sequential fashion based on statistical criterion that checks for the importance of variables. The importance of variables is defined in terms of a measure of the statistical significance of the coefficient for the variable. In stepwise selection procedure, backward and forward selection procedures are used simultaneously.

The final decision on the inclusion of each predictor variable will be made on the examination of the Wald statistic for the variable. Variables that do not contribute to the model based on these criteria would be eliminated and a new model should be fit. The new model would be compared with the old model through the LR test and AIC value. Also, the estimated coefficients for the remaining variables were compared to those from the full model. In view of this (deletion, refitting or verifying) was performed.

3.3.2.3 Model Adequacy Checking

Assessing goodness of fit involves investigating how close values are predicted by the model with that of observed values (Bewick *et al.,* 2005). Goodness of fit of the model can be assessed by overall model evaluation and testing the significance of each explanatory variable in the model. Clearly, the fitted model is good if there is a good agreement between the fitted and the observed data.

3.3.2.4 Statistical Tests of Individual Predictors

Wald Test

The Wald statistic is a test which is commonly used to test the significance of the individual logistic regression coefficients for each independent variable (that is, to test the null hypothesis in logistic regression that a particular coefficient is zero).

In logistic regression we have a binary outcome variable and one or more explanatory variables. For each explanatory variable in the model there will be an associated parameter. Therefore wald test, described by Agresti (1990), is used to test whether the parameter associated with an explanatory variable is zero or not. For a particular explanatory variable, or group of explanatory variables, if the Wald test is significant, then we would conclude that the parameters associated with these variables are non-zero, so that the variables should be included in the model (Agresti, 1990). For a dichotomous dependent variable the Wald test statistics is given by:

$$W = Z^2 = \frac{\hat{\beta}_i^{\,2}}{Var(\hat{\beta}_i)} \sim X^2_{(1)} \tag{3.12}$$

Under the null hypothesis $H_0: \beta_i = 0$ vs $H_1: \beta_i \neq 0$, for i = 1, 2,…,k. The statistics W is approximately distributed as chi-square with one degree of freedom.

3.3.2.5 Statistical Tests for Goodness Fit of Overall Model

Likelihood ratio test

The likelihood ratio test (LRT) is the most common test for assessment of overall goodness of fit for logistic regression model. The test statistic is defined as negative two times the natural logarithm of the ratio of likelihood functions of two models evaluated at their maximum likelihood estimates (MLEs). The likelihood-ratio test uses the ratio of the maximized value of the likelihood function for the full model(L_1) over the maximized value of the likelihood function for the reduced model (L_0). Therefore, the likelihood-ratio test statistic is given by:

$$G^2 = -2\ln\left(\frac{L_0}{L_1}\right) = -2(\ln L_0 - \ln L_1) \tag{3.10}$$

Where L_0 is the likelihood function of the reduced model and L_1 is the likelihood function of the full model evaluated at the maximum likelihood estimate. The LR test statistics G^2 is distributed chi-square with degrees of freedom equal to the difference between the numbers of parameters estimated in the two models (Menard, 2002). It is important to test the null hypothesis that all population logistic regressions coefficients are not significantly different from zero.

Hosmer-Lemeshow Test

The Hosmer-Lemeshow goodness-of-fit test is used to assess whether the number of predicted events from the logistic regression model reflect the number of observed events in the data. The data are ranked according to the predicted probability of the outcome from the model that is being evaluated (Hosmer and Lemeshow, 2000).

Hosmer-Lemeshow test is based on grouping cases in deciles in the sense that it is obtained by applying a chi-square test on a 2×g contingency table. The contingency table is constructed by cross classifying the dichotomous dependent variable with approximately g=10 groups in which the groups are formed by partitioning the predicted probabilities using the percentiles of the predicted event probability. It evaluates the goodness of fit by creating these 10 ordered groups of subjects and then compares in each observed group to the number predicted by the logistic regression model. The 10 ordered groups are created based on their estimated probability in such a way that those with estimated probability below 0.1 form one group, and so on, up to those with probability 0.9 to 1. Each of these categories is further divided into two groups based on the actual observed outcome variable (success and failure) (Hosmer and Lemeshow, 2000).

The expected frequencies for each of the cells are obtained from the model. If the model is good, most of subject with success are classified in the higher deciles of risk and those with failure in the lower deciles of risk and if the significance of the test is less than 0.05, then the model does not adequately fit the data. Thus, the test statistic is a chi-square statistic with a desirable outcome of non-significance, indicating that the model prediction does not significantly differ from the observed. The Hosmer and Lemeshow test statistic is given by:

$$\hat{C} = \sum_{j=1}^{g} \frac{\left(O_j - E_j\right)^2}{V_j} \tag{3.11}$$

Where, $E_j = n\pi_j$, $V_j = n\pi_j\left(1 - \pi_j\right)$, g is the number of group, O_j is the observed number of events in the j^{th} group, E_j is the expected number of events in the j^{th} group, V_j is the variance correction factors for the j^{th} group (Agresti, 2008). If the observed number of events differs from what is expected by the model, the statistic will be large and there will be evidence against the null hypothesis that the model is adequate to fit the data. This statistic has an approximate chi-square distribution with degree of freedom g-2 (Hosmer and Lemeshaw, 2000).

3.3.3 Multilevel Logistic Regression Model

Multilevel logistic regression model is appropriate for research designs where data for respondents are organized more than one level (i.e. nested data). Multilevel models have been developed to allow analysis at several levels simultaneously (hence the name multilevel), rather than having to choose at which level to carry out a single level analysis. This enable the extent of variation in the outcome of interest (in this case women early marriage) to be measured at each level assumed in the model both before and after the inclusion of explanatory variables in the model. In this research the individual women are nested in to regions so, two-level logistic regression model is a natural choice to use. In literature review multilevel logistic regression model also referred to as hierarchical model that can account for lack of independence across levels of nested data. Standards logistic regression assumes that all experimental units (in these case, married women) are independent in the sense that any variables affecting the dependent variable have the same effect in all regions. Multilevel modeling relaxes this assumption and allows these variables effects to vary across regions which can be used to analyze nested sources of variability in hierarchical data, taking in to account the variability associated with each level of the hierarchy (Snijders and Bosker, 1999)

In the present study, two-level binary logistic regression model was adopted. That means models accounting for married women-level and regional-level effects. The data structure, level-1 is the married women and level-2 is the regional level. Within each j^{th} region there are n_j individual married woman. Conceptually, the basic multilevel model for a binary response is equivalent to equation (3.5) except for the notation in the outcome variable.

In EDHS, 2016 we have data consisting of married woman nested into regions. Let Y_{ij} be the binary response for married woman i in region j and x_{ij} the associated factors of women early marriage. We define the probability of women early marriage as:

$\pi_{ij} = P(y_{ij} = 1)$ and let π_{ij} be modeled using a logit link function. The standard assumption is that y_{ij} has a Bernoulli distribution where the two-level model can be written as (Snijders and Bosker, 1999).

$$\mathrm{Logit}(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 X_{ij} + U_j \qquad (3.13)$$

Where, $U_j \sim \mathrm{IID}\ (0, \sigma^2{}_u)$, U_j is the random effect at level 2, without U_j, this equation can be considered as a standard logistic regression model. Therefore, conditional on U_j the Y_{ij} can be assumed to be independently distributed. In order to know the variation at each level, the equation (3.13) can be written by splitting up into two models: One for individual level and the other for regional level.

$$\mathrm{logit}(\pi_{ij}) = \ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \beta_1 X_{ij} \quad \text{Model of level one}$$

And $$\beta_{0j} = \beta_0 + U_j \quad \text{Model of level two} \qquad (3.14)$$

The intercept consists of two terms: a fixed component, β_0 and a group-specific component, random effect, U_j (Snijders and Bosker, 1999).

Heterogeneous Proportions

Heterogeneous proportion is the basic data structure of two-level logistic regression which is a collection of N groups, units at level-two (regions) and within region j (j=1, 2… N) a random sample of n_j married women units. The outcome variable is dichotomous and denoted by Y_{ij} ($i = 1, 2, \ldots, n_j$, $j = 1, 2, \ldots, N$) married women i nested in to region j and the total sample size is $W = \sum_{j=1}^{N} n_j$. If one does not take explanatory variables into account, the probability of success is assumed constant in each group (Snijders and Bosker, 1999). Let the early married women in group (region) j be denoted by π_j. The dichotomous outcome variable for the individual married women i in region j. Y_{ij} can be expressed as the sum of probability women early married in group (region) j, $E(Y_{ij}) = \pi_j$ plus some individual-dependent residual ε_{ij} that is:

$$Y_{ij} = \pi_j + \varepsilon_{ij}.$$

The residual term is assumed to have mean zero and variance, $\mathrm{var}(\varepsilon_{ij}) = \pi_j(1 - \pi_j)$. since the outcome variable is coded 0 and 1, the group sample average is the proportion of women being early married in group (region) j given by:

$$\hat{\pi}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} Y_{ij}$$

Where $\hat{\pi}_j$ is an estimate for the group-dependent probability, π_j. Similarly the overall sample average is the overall proportion of early marriage, $\hat{\pi}$. and is given as:

$$\hat{\pi} = \frac{1}{W}\sum_{j=1}^{N}\sum_{i=1}^{n_j} Y_{ij} \text{ , where W is total sample size}$$

Testing Heterogeneous Proportions

For the proper application of multilevel analysis the first logical step is to test heterogeneity of proportions between groups. To test whether there are indeed systematic differences between the groups, the well-known chi-square test for contingency table can be used. In this case the chi-square test statistic is:

$$X^2 = \sum_{j=1}^{N} n_j \frac{(\hat{\pi}_j - \hat{\pi})^2}{\hat{\pi}(1 - \hat{\pi})} \tag{3.15}$$

This statistic follows approximately a chi-square distribution with $N - 1$ degrees of freedom.

Estimation of between and Within Groups variance

This is the true variance between regions dependent probabilities, i.e. the population values of $\mathrm{var}(\pi_j)$ can be estimated by:

$$\psi^2 = S^2_{between} - \frac{S^2_{within}}{\tilde{n}} \tag{3.16}$$

Where $\tilde{n} = \frac{1}{N-1}\left(w - \frac{\sum_{j=1}^{N} {n_j}^2}{W}\right)$

For dichotomous outcome variables, the observed between regions variance is closely related to the chi-square test statistic given in equation (3.15) (Snijders and Bosker, 1999). The between-groups (region) variance is given as:

$$S^2{}_{\text{between}} = \frac{\hat{\pi}(1-\hat{\pi})}{\tilde{n}(N-1)} X^2 \tag{3.17}$$

Where X^2 is as given by equation (3.15) and the within group variance in case of a dichotomous outcome variable is a function of group averages, by within-groups variance:

$$S^2{}_{\text{within}} = \frac{1}{W-N} \sum_{j=1}^{N} n_j \, \pi_j \left(1 - \pi_j\right) \tag{3.18}$$

i. Random Intercept only Multilevel Model

The empty two-level model for a binary outcome variable refers to a population of groups (i.e. regions) and specifies the probability distribution for group-dependent probabilities without considering further explanatory variables in to account. This model only contains random groups and random variation within groups (regions). It can be expressed with logit link function as follows (Snijders and Bosker, 1999).

$$\text{Logit}\left(\pi_j\right) = \beta_0 + U_{0j} \tag{3.19}$$

$$\text{Where,} \quad U_{0j} \sim \text{IID}\ (0, \sigma^2{}_0)$$

Where β_0 the population average of the transformed probability and U_{0j} is the random deviation from this average for group j. For the deviations U_{0j} it is assumed that they are independent random variables with a normal distribution with mean zero and variance $\sigma_0{}^2$. This model does not include a separate parameter for the individual level variance (Snijders and Bosker, 1999). This is because the individual level residual variance of the Y_{ij} (married women) follows Bernoulli distribution directly from the probability of having women early married (π_j) which is given by: $\text{var}(\boldsymbol{\varepsilon}_{ij}) = \pi_j(1 - \pi_j)$ denoted by π_0. Here the probability corresponding to the average value β_0, as defined by: $f(\pi_0) = \beta_0$.

ii. The Random Intercept and Fixed Slope Multilevel Model

In the random intercept model the intercept is the only random effect meaning that the groups differ with respect to the average value of the response variable, but the relation between explanatory and response variables cannot differ between groups. We assume that there are variables which potentially explain the observed success and failure. These variables are denoted by $X_h (h = 1, 2, \dots, k)$ with their values indicated by X_{hij}. Since some or all of these variables could be factors of level one (married women), the probability of women early marriage is not necessarily the same for all individual in a given group (Snijders and Bosker, 1999). Therefore, the success probability depends on the married women as well as the group (region), and is denoted by π_{ij}. The outcome variable is split into an expected value and residual as:

$$Y_{ij} = \pi_{ij} + R_{ij},$$

Where Y_{ij} is the dichotomous outcome variable for the individual married women i in region j, π_{ij} is the probability of i women early marriage in region j and R_{ij} is the residuals.

The random intercept model expresses the log-odds, i.e. the logit of π_{ij} as a sum of a linear function of the explanatory variables. That is (Snijders and Bosker, 1999).

$$\operatorname{logit}(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \cdots + \beta_k X_{kij}$$

$$= \beta_0 + \sum_{h=1}^{k} \beta_h X_{hij} + U_{0j} \qquad 3.20$$

Where, $\operatorname{logit}(\pi_{ij})$ does not include a married woman residual because it is an equation for the probability of having early married women π_{ij} rather than for the outcome Y_{ij}.

β_{0j} - is assumed to vary randomly and is given by the sum of an average intercept β_0 and group (region) dependent deviations U_{0j} is given:

$$\beta_{0j} = \beta_0 + U_{0j} \qquad (3.21)$$

The first part of equation (3.22) incorporating the regression coefficients $\beta_0 + \sum_{h=1}^{k} \beta_h X_{hij}$ is the fixed part of the model, because coefficients are fixed. The remaining part U_{0j} is called the random part of the model. It is assumed that the residual, U_{0j} are mutually independent and normally distributed with mean zero and variance $\sigma^2{}_0$ (Snijders and Bosker, 1999). As a result from Eq. (3.22) solving for π_{ij}.

$$\pi_{ij} = \frac{\operatorname{Exp}\left(\beta_0 + \sum_{h=1}^{k} \beta_h X_{hij} + U_{0j}\right)}{1 + \operatorname{Exp}\left(\beta_0 + \sum_{h=1}^{k} \beta_h X_{hij} + U_{0j}\right)} \qquad (3.22)$$

Where, β_h- is a unit difference between the X_h values of two individuals in the same group which associated with a difference of β_h in their log-odds, or equivalently, a ratio of $\exp(\beta_h)$ in their odds.

iii. The Random Coefficient Multilevel Logistic Regression Model

So far, we have allowed the probability of women early marriage to vary across regions, but we have assumed that the effects of the explanatory variables are the same for each region. Now we modify this assumption by allowing the difference between the effects of explanatory variables to vary across regions. To allow for this effect, we will need to introduce a random coefficient for those explanatory variables. So, a random coefficient model represents heterogeneity in relationship between the explanatory variables and the observed outcomes. Suppose that there are k factors of women early marriage $X_{1ij}, X_{2ij}, \dots, X_{kij}$, and consider the model where all predictor variables have varying slopes and random intercept. That is

$$\mathrm{Logit}(\pi_{ij}) = \mathrm{Log}\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \sum_{h=1}^{K} \beta_h X_{hij} + \sum_{h=1}^{k} U_{hj} X_{hij}$$

Letting $\beta_{0j} = \beta_0 + U_{0j}$ and $\beta_{hj} = \beta_h + U_{hj}$ where (h = 1, 2,…, k) we have:

$$\mathrm{Logit}(\pi_{ij}) = \mathrm{Log}\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \sum_{h=1}^{K} \beta_h X_{hij} + U_{0j} + \sum_{h=1}^{k} U_{hj} X_{hij} \qquad (3.23)$$

The first parts $\beta_0 + \sum_{h=1}^{K} \beta_h X_{hij}$ are fixed part of the model and the second parts $U_{0j} + \sum_{h=1}^{k} U_{hj} X_{hij}$ are called the random part of the model. The random variables or effects $U_{0j}, U_{1j}, U_{2j}, \dots, U_{kj}$ are assumed to be independent between groups but may be correlated within groups. So the components of the vector $U_{0j}, U_{1j}, U_{2j}, \dots, U_{kj}$ are independently distributed as a multivariate normal distribution with zero mean vector and variances and co-variances matrix given by:

$$\omega = \begin{bmatrix} {\sigma_0}^2 & . & \cdots & . \\ \sigma_{10} & {\sigma_1}^2 & \cdots & . \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k0} & \sigma_{k1} & \cdots & {\sigma_k}^2 \end{bmatrix}$$

Intra-class Correlation Coefficient (ICC)

The other fundamental reason for applying multilevel analysis is the existence of intra-class (intra-regional) correlation arising from similarity of early marriage for women in the same region compared to those of different regions. The intra-class correlation coefficient (ICC) measures the proportion of variance in the outcome explained by the grouping structure. ICC can be calculated using an intercept-only model. This model can be derived from equation [3.23] by excluding all explanatory variables, which results in the following equation (snijder and Bosker, 1999):

$\text{logit}(\pi_j) = \beta_0 + U_{0j},$ then ICC is calculated based on the following formula:

$$\text{ICC} = \frac{\sigma^2{}_{uo}}{\sigma^2{}_{uo} + \sigma^2{}_e}$$

Where, $\sigma^2{}_e$ is variance of individual (lower) level units. In multilevel logit model level one residual is $\sigma^2{}_e = \pi^2/3 \approx 3.29$ (snijder and Bosker, 1999) and this formula can be reformulated:

$$\text{ICC} = \frac{\sigma^2{}_{u_0}}{\sigma^2{}_{u_0} + 3.29} \qquad (3.24)$$

3.3.3.1 Parameter Estimation for Multilevel Model

Like the methods for ordinary logistic regression, Parameter estimation for multilevel logistic model is not straightforward. The most common methods to estimate parameters in multilevel logistic regression model were based on likelihood (ML) method. This method has two prevailing approximation, marginal quasi-likelihood (MQL) (Goldstein, 1991; Goldstein and Rasbash; 1996) and penalized quasi-likelihood (PQL) (Laird, 1978; Breslow and Clayton, 1993). After applying these quasi likelihood methods, the model will be estimated using iterative generalized least squares (IGLS) or reweighted IGLS (RIGLS) (Goldstein, 2003). Both MQL and PQL are based on Taylor series expansion to achieve the approximation of parameter estimation. Based on the first and second term of Taylor series expansion, MQL and PQL are often known as first order and second order of MQL and PQL, respectively. However the maximum likelihood method parameter estimation using penalized quasi-likelihood doesn't provide the model comparison statistics (deviance and information criteria) while marginal quasi-likelihood (MQL) using approximation of the integrand Laplace approximation and Gausse-Hermite quadrature using R software can provide. In addition to that it is more complicated due to the need to perform numerical integration to obtain a marginal likelihood to maximize. So that we used approaches based on Laplace approximation and Gausse-Hermite quadrature. In glmer function of R software it is possible to specify number of Gauss–Hermite quadrature points only when there is one random effect (i.e., for the random intercept model). When there are two or more random effects (e.g., random intercept and slope), due to greater computational burden, only a single number of adaptive quadrature point (i.e., Laplace approximation) is allowed. This is the limitation of glmer function in generalized linear mixed model. Therefore in this study the researcher forced to use Laplace approximation of the maximum likelihood method to estimate parameters.

3.3.3.2 Multilevel Model Comparison

Likelihood Ratio Test

It is often necessary to see the goodness of fits of different models. When fitting several models to the same data set, likelihood ratio test can be helpful to the goodness of fit using the deviance based chi-square test. The likelihood ratio test compares the deviance (-2 log likelihood) of two models by subtracting the smaller deviance (model with more parameters) from the larger deviance (model with reduced number of parameter). The difference is a chi-square test with the number of degrees of freedom equal to the difference number of different parameters in the two models. In the case where the empty model is compared to a full model, the likelihood ratio test provides information about whether the predictors in the model together account for a significant amount of variance in the dependent variable (Agresti, 2002).

Information criteria's

AIC and BIC also used at assessing the model goodness of fit. The AIC and BIC fit indices are based on the deviance statistic, but they incorporate penalties for a greater number of parameters. The smaller the AIC or BIC value, the better is the model. These are defined as:

$$\mathrm{AIC} = -2\log(\mathrm{Likelihood}) + 2\mathrm{K} \quad (3.25)$$

$$\mathrm{BIC} = -2\log(\mathrm{Likelihood}) + \mathrm{K} * \log(\mathrm{N}) \quad (3.26)$$

Where k is the model degrees of freedom calculated as the rank of variance–covariance matrix of the parameters and N is the number of observations used in estimation. AIC and BIC can be viewed as measures that combine fit and complexity. Fit is measured negatively by -2*ln (Likelihood). The larger the value, the worse the fit is. Complexity is measured positively, either by 2*k(AIC) or ln(N)*k(BIC). The larger the value also, the worse the fit is. (Akaike, 1974).

4. RESULTS AND DISCUSSIONS

The aim of this chapter is to describe and make analysis to investigate factors that influence women early marriage in Ethiopia based on EDHS, 2016 data. The nature of data set was based on two-stage stratified cluster sampling in which the lower levels (married women) are nested within units at higher level (region).

4.1 Summaries of Descriptive Statistics

Descriptive analysis is a process of describing data set by tables, graphs and summary calculations. In this study, the researchers employed cross tabulation to describe both dependent and independent variables.

A total of 9825 married women from nine regional states and two administrative cities of Ethiopia were considered for the analysis. Among a total of women participated in this study about 5976 (60.8%) married at early age while 3849 (39.2%) married at an age of 18 years and above. Based on Table 4.1 the percentage of early marriage varied based on various demographic and socioeconomic factors. The data on demographic and socioeconomic factors characteristics of early marriage were displayed in Table 4.1. As it can be observed, from a total higher number of women early marriage 4838 (80.96%) was recorded in rural areas, and relatively small number of women early marriage 1138 (19.04) recorded in urban areas.

The percentage of women got married at early age varied from one region to another in Ethiopia. The highest proportion 835(13.97%) and 833(13.94%) of women early marriage were observed in Amhara and Oromia respectively which almost equal and followed by SNNP 695(11.63%) whereas the least proportion of early marriage was recorded in Addis Ababa 216(3.61%) and followed by Dire Dawa 314(5.25%). Hence, it is an indication for variation in the proportion of women early marriage among region of Ethiopia.

Table 4.1 also revealed that the number of women early marriage was varied by religion. The highest percentage 2771(46.37%) was recorded in Muslim religion followers followed by orthodox believers 2100 (35.14%). On the other hand, protestant women tend to occupy an intermediate position in early marriage 1035 (17.32%) and the other religion take the least position in number of early marriage 70 (1.17%). The number of women early marriage was also varied by level of education. Thus, the percentage of women early marriage is 3751 (62.77%) for non educated women, 1579 (26.42%) for primary educated women, 434(7.26%) for secondary educated women and 212 (3.55%) for women whose their education level is higher.

It was reported from the Table 4.1 that women came from the poorest family have high chance of early marriage which account 1933 (32.35%). Similarly, the number of early marriage for women from richer 858 (14.36%) and middle 907 (15.18%) parent wealth index. With regard to women work status, the percentage of unemployed women had higher chance to be early married 4039 (67.59%) than women who are employed 1937 (32.51%).

Table 4.1 Distribution of Socio-economic and Demographic related determinant factors of women early marriage in Ethiopia (EDHS, 2016).

variable	Categories	Age at First Marriage			Chi-square	P-value
		Legal Marriage n(%)	Early marriage n (%)	Total n (%)		
Region	Tigray	341 (8.86)	620(10.37)	961 (9.78)	436.468	1.6e-87
	Affar	221 (5.74)	637 (10.66)	858 (8.73)		
	Amhara	294 (7.64)	835 (13.97)	1129 (11.49)		
	Oromia	488 (12.68)	833 (13.94)	1321 (13.45)		
	Sumale	419(10.89)	556 (9.30)	975 (9.92)		
	B/Gumes	292 (7.59)	511 (8.55)	803 (8.17)		
	SNNPR	515 (13.38)	695 (11.63)	1210 (12.32)		
	Gambella	270 (7.01)	443 (7.41)	713 (7.26)		
	Harari	259 (6.73)	316 (5.29)	575 (5.85)		
	A/Ababa	463 (12.03)	216 (3.61)	679 (6.91)		
	D/Dawa	287 (7.46)	314 (5.25)	601(6.12)		
Place of Residence	Urban	1370 (35.59)	1138 (19.04)	2508 (25.53)	337.3414	3.7e-75
	Rural	2479 (64.41)	4838 (80.96)	7317 (74.47)		
Religion of Women	Orthodox	1488 (38.66)	2100 (35.14)	3588 (36.52)	53.1194	1.7e-11
	Muslim	1596 (41.47)	2771 (46.37)	4367 (44.45)		
	Protestant	756 (19.64)	1035 (17.32)	1791 (18.23)		
	Others	9 (0.23)	70 (1.17)	79 (0.80)		
Women Education	None	1,961 (50.95)	3,751 (62.77)	5712 (58.14)	188.736	1.1e-40
	Primary	1164 (30.24)	1579 (26.42)	2743 (27.92)		
	Secondary	420 (10.91)	434 (7.26)	854 (8.69)		
	Higher	304 (7.90)	212 (3.55)	516 (5.25)		
	Richest	1455 (37.80)	1284 (21.49)	2739 (27.88)	311.926	2.8e-66
	Richer	459 (11.93)	858 (14.36)	1317 (13.40)		
	Middle	460 (11.95)	907 (15.18)	1367 (13.91)		

Wealth Index	Poorer Poorest	476 (12.37) 1000 (25.98)	993 (16.62) 1933 (32.35)	1469 (14.95) 2933 (29.85)		
Women work	Working Nonworking	1435 (37.28) 2414 (62.72)	1937 (32.51) 4039 (67.59)	3372 (34.32) 6453 (65.68)	24.6262	7.7e-07
Media Exposure	Yes No	1204 (31.28) 2645 (68.72)	1285 (21.50) 4691 (78.50)	2489 (25.33) 7336 (74.67)	117.3375	1.8e-27
Husband Education level	None Primary Secondary Higher	1518 (39.44) 1128 (29.31) 596 (15.48) 607 (15.77)	2985 (49.95) 1940 (32.46) 634 (10.61) 417 (6.98)	4503 (45.83) 3068 (31.23) 1230 (12.52) 1024 (10.42)	282.0071	7.7e-61
Husband Occupation	Agriculture Professional Business Labourers Others	1359 (35.31) 1003 (26.06) 606 (15.74) 36 (0.94) 845 (21.95)	2780 (46.52) 1374 (22.99) 639 (10.69) 32 (0.54) 1151 (19.26)	4139 (42.13) 2377 (24.19 1245 (12.67) 68 (0.69) 1996 (20.32)	139.8685	3.0e-29
No. of sibling	< = 5 >5	2019 (52.46) 1830 (47.54)	2931 (49.05) 3045 (50.95)	4950 (50.38) 4875 (49.62)	10.8834	0.001

Table 4.1 shows that the percentage of early marriage for women not exposed to any mass media messages via Radio, TV or newspapers were 4691(78.50%) as compared to exposed group 1285(21.50%). With regard to husband's education level, women whose their husbands not educated were highly early married 2985(49.95%) than women with primary 1940 (32.46%), secondary 634 (10.61%) and higher 417 (6.98%) educated husband's.

As reported on the Table 4.1 in Ethiopia the proportion of women early marriage depends on husband's occupation. The highest proportion was observed among women who's their husband/partners occupation was agriculture 2780 (46.52%) while the lowest proportion of women early marriage was recorded in women who's their husband occupation was laborers 32 (0.54%). Relating total number of sibling 3045 (50.95%) early marriage were observed among women came from family size greater than five while, 2931 (49.05%) were occurred among women came from parents with family size less than five, which was almost equal.

From Table 4.1 there were also some inferential statistics which are used for test of association. Thus, the chi-square test was carried out to determine the association between the marriage stage among women and the independent variables (region, place of residence, religion of respondent, women education level, wealth index, women work status, media exposure, husband education level, husband occupation status and total number of sibling). The result revealed that all the included independent variable had statistically significant association with women early marriage at 0.25 level of significant. Binary logistic regression model prediction for women early marriage using EDHS, 2016 is analyzed based on those variables has significant association with women early marriage stated earlier. Hence, we should identify statistically significant predictor variables and determine the direction of relationship with the dependent and independent variables using single level and multilevel logistic regression analyses.

4.2 Associated Factors of Women Early Marriage in Ethiopia

The results of logistic regression analysis were obtained by using stepwise inclusion of variables that has significant association with women early marriage. Thus, the overall model evaluation, statistical tests of individual predictors and goodness-of-fit statistics are presented. In binary logistic regression analysis the initial log likelihood function, (-2 Log Likelihood) before any variable take in to account is fitted. If the associated factors have a relationship with women early marriage, we will improve our ability to predict the dependent variable accurately, and the log likelihood value will decrease. Thus, the initial –2LL value is 13156.2 at step 0, before any variables was added to the model. Finally, all associated factors of women early marriage are added to the logistic regression equation in a stepwise manner. The addition of these variables reduced the initial log likelihood value (-2Log Likelihood) of 13156.2 to 12465.45. The statistical significance of individual regression coefficients were tested using the Wald and score chi-square statistic. Therefore the result presented in Table 4.7 on Appendix showed that region, place of residence, religion of respondent, women education level, wealth index, husband education level, husband occupation status and total number of sibling were found to be statistically significant predictors of women early marriage at 5% level of significance.

The result of binary logistic regression shows statistically significant variables, the direction of relationship and the maximum likelihood estimates of the parameters. A negative sign in column labeled estimate (see Appendix: Table 4.7) indicates that the effects of the category on the log odds of the response variable is less likely appear as compared to the reference category. In contrast a positive coefficient column labeled estimate indicates that the effects of the category on the log odds of the response variable appear more likely as compared to the reference category. A more appealing way to interpret the regression coefficient in logistic model is using odds ratio. The odds ratio indicates the effect of each associated factors of women early marriage directly on the odds of women early marriage rather than on the odds of legal marriage. Estimates of OR greater than one indicate that the proportion of women early marriage is greater than that for the reference category while OR less than one shows the proportion of women early marriage is less than that for the reference category of each variable.

4.2.1 Goodness Fit of Logistic Regression Model

For categorical data, after a logistic regression model has been fitted, a global test of goodness of fit of the resulting model should be performed. It is necessary to see the appropriateness, adequacy and usefulness of the fitted model. The most commonly used techniques are Likelihood-Ratio test, Hosmer and Lemeshow test and the Wald test goodness of fit.

Likelihood-Ratio Test:- is the most common assessment of overall model fit for logistic regression is the likelihood ratio test, which is the deviance chi-square difference between the null model and model containing a set of predictors. The difference between -2log-likelihood for model fitted with independent variables and -2log-likelihood for null model (at step 0, before any variables have been added to the model) is distributed $\chi 2$ with degrees of freedom equal the difference between the numbers of parameters in the two models.

Table 4.2 Summary of Model Fit Statistics for Intercept only and Full Model

Goodness of Fit Measure	Empty Model	Full Model
-2LL	13156.2	12465.45
AIC	13158.2	12529.62
BIC	13165.4	12759.62

As shown in Table 4.2 the -2log-likelihood value for the null (intercept only) model and full model were 13,156.2 and 12,465.45, respectively. To test the significance of the full model over the null model the likelihood ratio test provided a chi-square value of 690.5 ($p<0.0001$) which would imply good fit for the full model. Moreover, the model with the smallest values of AIC and BIC is also considered as the best fit. Thus, the value of AIC and BIC for null model was 13,158.2 and 13,165.4, and for full model was 12,529.62 and 12,759.62, respectively. Hence, the full model fits the data well, indicating that the included associated factors of women early marriage had significant effect with the dependent variable.

Hosmer-LemeshowTest:- is a test of assessing goodness fit of the model. The Hosmere-lemeshow test was applied and the result presented in Table 4.3. Well-fitting model is indicated by an insignificant chi-square value and confirming that there is no difference between the observed and the model predicted values. Since the calculated chi-square 11.030 is less than the tabulated chi-square 15.507 at 8 degree of freedom and the p-value 0.200 was greater than 0.05 sig level, the model is good fit.

Table 4.3 Test of Significance of Hosmer-Lemeshow Goodness of Fit Statistics

Chi-square	DF	P-value
11.030	8	0.200

4.3 Results of Multilevel Logistic Regression Analysis

The data used in this study have a hierarchical structure. Units at one level are nested within units at the next higher level. Here, the lower level (level-one) units are married women, and the higher level (level-two) units are groups (regions). The nesting structure is married women within regions that resulted in a set of 11 regions with a total of 9825 married women. Also because of clustering effect of the higher (regional) level there was loss of independence among individual married women which was accounted using multilevel logistic regression analysis.

Testing Heterogeneity Proportions of Women Early Married Among Regions of Ethiopia

The first step in performing a multilevel analysis is testing the heterogeneity of proportions between groups (regions). Chi-square test was applied to assess heterogeneity in the proportion of women early marriage among regions. The test yields the calculated $X^2 = 436.468$ with $P = 0.001 < 0.05$. Thus, there is an evidence of heterogeneity of women early marriage among regional state of Ethiopia. Therefore, multilevel logistic regression model was attempted. As we did in the single level binary logistic regression model, we should do model comparison and model selection among the three candidate multilevel logistic regression models before interpreting the coefficients.

4.3.1 Comparison of Multilevel Logistic Regression Models

Before interpreting the results of multilevel model analysis, we compare the three candidate multilevel logistic regression models (nested models) which should be based on the necessity of parsimony in a model. To do so, LRT (deviance based chi-square) test and information criteria/model diagnostic statistic (AIC and BIC) were used to select the best fitted model among the three fitted two-level logistic regression models.

The deviance-based chi-square value ($\chi 2 = 383.94$, d.f = 1, p-value = $0.001 < 0.05$) for the empty model is shown in Table 4.3. This deviance-based chi-square statistics is calculated as the difference of log likelihoods between an empty model for single level logistic regression (Table 4.3) and random intercept only multilevel logistic regression model (Table 4.4), which is to be compared with the critical value from the chi-square distribution with 1 degree of freedom. The significance of this test further implies that an empty model with random intercept is more appropriate than an empty model without random intercept to fit the data.

Similarly, the deviance based on chi–square test statistics for multilevel random intercept with fixed slope logistic model and multilevel random intercept only model shown in Table 4.4. It is compared with the critical value from the chi-square distribution with 21 degree of freedom (i.e. $\chi 2 = 260.85$, d.f =23-2 = 21, p-value = $0.001 < 0.05$). This suggesting that random intercept with fixed coefficient multilevel model fits better the data set as compared to the random intercept only model.

The deviance based on chi-square value for multilevel random coefficient binary logistic regression and random intercept model also shown in Table 4.4. The statistics was similarly compared with the critical value from the chi-square distribution with 5 degree of freedom $\chi 2 = 5.74$, d.f =28-23 = 5. Since p-value = 0.331 which is greater than 5% level of significance suggesting no evidence against the significance of random coefficient multilevel logistic regression model. Furthermore, the information criteria (AIC and BIC) were also used to make an overall model adequacy comparison of the three multilevel model. Accordingly, Table 4.4 reveals that the computed AIC and BIC value for the random intercept with fixed coefficient model (AIC = 12,557.39 and BIC = 12,722.82) were less than that of the random coefficient model (AIC = 12,561.66 and BIC = 12,763.05) and the random intercept only model (AIC = 12,776.26 and BIC = 12790.64) respectively. This indicated that the random intercept with fixed coefficient model was a better fit as compared to the random intercept only and the random coefficient multilevel logistic regression model.

Table 4.4 Comparison of multilevel logistics models using Information criteria and LRT

Model comparison statistics	Random intercept Only multilevel model	Random intercept with fixed coefficient model	Random coefficient multilevel model
-2*Log-likelihood	12,771	12,511.39	12,505.65
Deviance based $\chi 2$	283.946	260.8586	5.734
Degree freedom	2	23	28
p-value	0.000	0.000	0.3319
AIC	12,775	12,557.39	12,561.66
BIC	12,789	12,722.82	12,763.05

Therefore, based on the summary results of deviance and information criteria (AIC and BIC) (See Table 4.4) we conclude that random intercept multilevel model is better than other multilevel logistic regression model in predicting early marriage among women in Ethiopia.

4.3.2 Results of Random Intercept only Model

We first fitted an empty model with no explanatory variable (intercept-only model) that predicts the probability of women being early married. The simplest non-trivial specification of the hierarchical linear model is a model in which only the intercept varies between level two units. The empty model contains no explanatory variables and it can be considered as a parametric version of assessing heterogeneity among regions with respect to the prevalence of women early marriage.

Table 4.5 Results of Parameter Estimate of Intercept-Only Model with Random Effect.

Fixed effect	Estimate	S.E	Z	P-value	OR	[95% Conf. Interval]	
Intercept (β_0)	0.4013	0.142	2.826	0.0047	1.493	0.123	0.679

Random-effects	Estimate	Standard Error	P-value	[95% Conf. Interval]	
Region,$\mathrm{var}(U_{0j}) = \sigma^2{}_{u0}$	0 .2163	0.09488	0.01131	0.09161	0.51108

AIC = 12775, BIC = 12789, -2*Log-likelihood = 12,771

The intercept β_0 also known as the grand mean that is shared by all regions. As shown in Table 4.5, the log-odds of women early marriage given in all regions under investigation was estimated as $\beta_0 = 0.4013$. Table 4.5 also includes the variance estimate of random effects at regional level, $\sigma^2_{u0} = 0.2163$. The random effects of intercept implies that the between region variance of women early marriage is 0.2163 and tells that there is a significant difference in proportion of women early marriage across regions of Ethiopia. This statistical significance of random effect was tested as follows.

The random effect test examine hypothesis that whether or not the random intercept or between-region variance is needed for these data or not and this is statistically stated as:

$$H_0: \sigma^2_{u0} = 0 \quad \text{Versus} \quad H_1: \sigma^2_{u0} > 0$$

Variance test is not similar with the classical testing of other parameters. Because the constrained variance component test lies on the boundary of the parameter space, the likelihood ratio test can break down asymptotically. It has been shown that tests for variance component can be carried out using mixtures of chi-square distributions. In this study, we show that the null distribution of this one sided likelihood ratio test statistic converges to a 50:50 mixture of chi-square distributions with 0 and 1 degree of freedom given as:

$$0.5X_0^2+0.5X_1^2$$

$$\text{p} - \text{value} = 0.5\Pr(X^2_{0:1} > \text{Likelihood ratio test})$$

$$= 0.5\Pr(X^2_0 > 383.946) + 0.5\Pr(X^2_1 > 383.946) = 0.001$$

The likelihood ratio test statistic is equal to twice the difference of the log likelihoods of null model without random effect and the log likelihoods of null model with random effect model, or the difference of likelihood under null hypothesis and alternative hypothesis respectively. This is stated as:

$$\text{LRT} = -2\{l(Y|H_0) - l(Y|H_1)\}$$

Where $l(Y|H_0)$ and $l(Y|H_1)$ are the log likelihoods under the null and alternative hypotheses evaluated at their maximum likelihood estimates, respectively. The critical value for an $\alpha = 0.05$ test using this mixture distribution is 3.84, indicating we would reject H_0 for p-value = 0.001 <

0.05. This indicated that adding region as a random effect or between-region variance in the model was necessary in order to assess variation of women early marriage among regions. This confirming that there was statistically significance difference in proportion of women early marriage across regions providing that multilevel model is the best option to account the regional variation of women early marriage.

The empty model with random intercept also helps to calculate the between region variations by the help of intra–class correlation coefficient (ICC) which is the measure of the correlation between two individuals who are in the same higher level unit (region). ICC for this model is calculated by using formula (3.24). A low ICC indicates that relatively small between region variations. From table 4.5 the variance between regions was 0.2163 and variation among individual married women was $\frac{\pi^2}{3} = 3.29$. So, Intra-class correlation coefficient is 0.0616. This implied that about 6.16% of the variation in women early marriage can be explained by grouping the married women in to regions .The remaining (100 - 6.16% = 93.84%) of the variation in experiencing early marriage was explained within region-lower level units.

Now we examine estimates of the region effects or residuals $\widehat{U}_{0j}$ obtained from the null model. To calculate the residuals and produce a caterpillar plot with the region effects shown in rank order together with 95% confidence intervals. The plot 4.1 shows the estimated residuals for all 11 regions in the sample. The residuals represent regional departures from the overall mean, so a region whose confidence interval does not overlap the line at zero (representing the mean value of early marriage practice across all regions) is said to differ significantly from the average at the 5% significance level. At the lower side of the plot, there is a cluster of regions whose mean practice of early marriage was lower than average and vice versa.

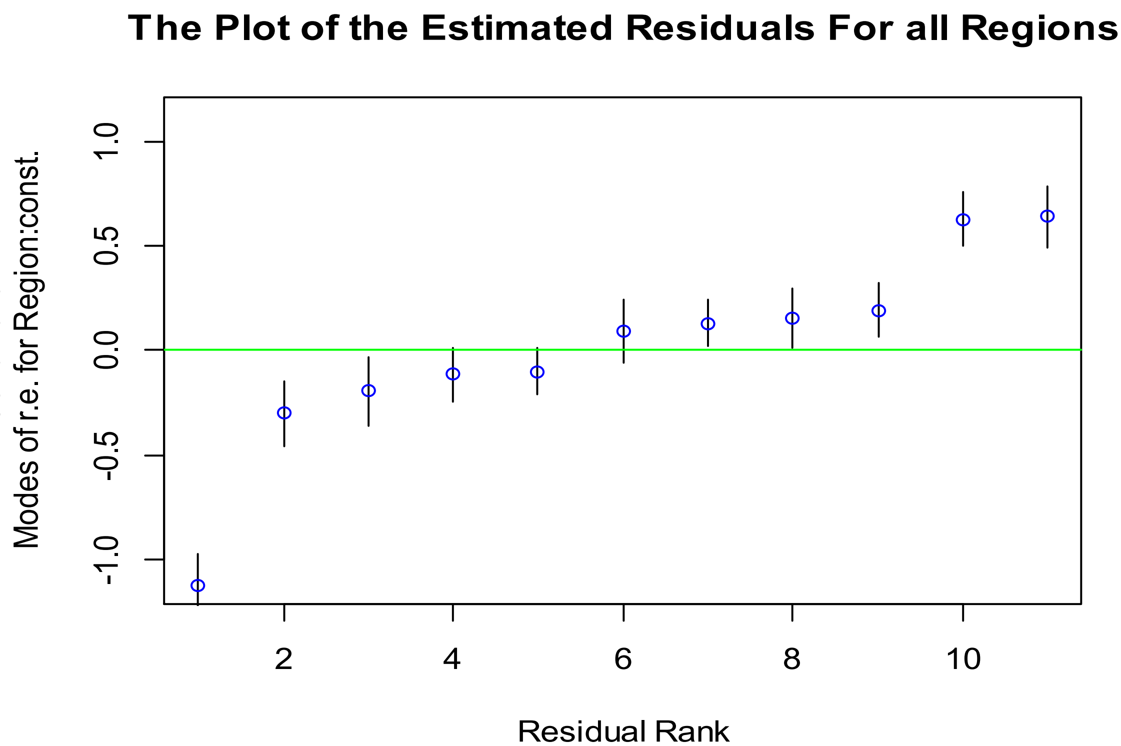


Figure 4.1 plot of estimated residuals (random effects) for all regions of women early marriage.

4.3.3 Results of Random Intercept Multilevel Analysis

To assess the effect of independent variables on women early marriage, we considered random intercept model. That is the probability of women early marriage is vary across regions, but we assumed that the effects of each explanatory variables are the same for each region. As shown in Table 4.5 the variance component of random intercept only multilevel model is 0.2163, whereas the variance for the random intercept with fixed coefficient model is 0.1045. The variance of random intercept logistic model decreased compared to that of the random intercept only model. The reduction of the random effect of the intercept variance is due to the inclusion of fixed explanatory variables. That is, taking in to account the fixed independent variables can provide extra predictive value on women early marriage among region.

Table 4.6 Result of Parameter Estimation for random intercept with fixed coefficient multilevel logistic regression model.

Fixed effect parameter estimated

Covariates	Categories	Estimate	Std. Err.	P-value	OR	[95%C.I for OR]
Intercept	constant	0.28702	0.13692	0.03614**	1.332	1.018, 1.742
Place of Residence	Urban(ref.)					
	Rural	0.2148	0.090	0.017 *	1.239	1.038, 1.479
Religion of Women	Orthodox(ref.)					
	Muslim	0.1629	0.065	0.012 **	1.176	1.035, 1.337
	Protestant	0.0343	0.077	0.658	1.034	0.888, 1.205
	Others	1.6933	0.361	<0.001***	5.437	2.679, 11.036
Women Education	None (ref.)					
	Primary	-0.1383	0.053	0.009*	0.870	0.784, 0.966
	Secondary	-0.0728	0.087	0.404	0.929	0.783, 1.103
	Higher	-0.236	0.112	0.035*	0.789	0.633, 0.983
Wealth Index	Richest(ref.)					
	Richer	0.1964	0.098	0.046*	1.217	1.002, 1.477
	Middle	0.2192	0.101	0.030*	1.245	1.020, 1.518
	Poorer	0.2816	0.101	0.005*	1.325	1.086, 1.616
	Poorest	0 .1140	0.097	0.243	1.120	0.925, 1.357
Women work	Working (ref.)					
	Not working	-0.0180	0.047	0.704	0.982	0.894, 1.078
Media Exposure	Yes(ref.)					
	No	-0.0941	0.062	0.130	0.910	0.805, 1.028
Husband education level	None (ref.)					
	Primary	0.0617	0.053	0.246	1.063	0.958, 1.180
	Secondary	-0.1647	0.074	0.027 **	0.848	0.732. 0.981
	Higher	-0.5411	-0.086	<0.001***	0.582	0.491, 0.689
Husband Occupation	Agriculture(ref.)					
	Professional	-0.0650	0.062	0.295	0.937	0.829, 1.058
	Business	-0.2246	0.073	0.002***	0.798	0.691, 0.923
	Labourers	-0.3037	0.255	0.234	0.738	0.447, 1.217
	Others	-0.1579	0.061	0.010 **	0.853	0.756, 0.963
No. of sibling	< = 5 (ref.)					
	>5	0.1020	0.043	0.019*	1.107	1.016, 1.205
Random effect (Region)		Variance component		S.E	[95% Conf.I] for σ^2_{u0}	
Between Region variance, σ^2_{u0}		0.1045166		0.047	0.0427, 0.2557	

In the results of the random intercept with fixed slope model, the fixed part showed that place of residence, religion of respondent, women education attainment, wealth index, husband education attainment, husband occupation status and total number of sibling were found to be statistically significant factors of women early marriage in Ethiopia at 5% level of significance.

Interpretation of Results from Random Intercept Multilevel Logistic Regression Model

Table 4.6 presents parameter estimates and their corresponding empirically corrected standard errors alongside the p-values from random intercept with fixed slope multilevel model. Each parameter β_j reflects the effect of factor X_j on the log odds of the probability of women early marriage statistically controlling all the other covariates and random effects in the model. Then, the odds ratio of variables is calculated as the exponent of β_j i.e. $OR = \exp(\hat{\beta})$.

This model revealed that the odds of early marriage, women who lived rural area were 1.239 (OR=1.239, CI: 1.038, 1.479) times those lived in urban.

The estimated odds of Muslim religion was 1.176, (CI: 1.035, 1.337). This means Muslim believer's women were 17.6% more experience early marriage than those orthodox believers. Similarly the prevalence of women early marriage was highly occurred with women believes in other religion (excluding orthodox, Muslim and protestant) as compared to women orthodox believers. The odds ratio of women early marriage in other religion is 5.437(CI: 2.679, 11.036) times that of Orthodox believers. This implies that the probability of women early marriage is 5.437 times more likely for women who believe other religion than orthodox believers. Statistically significant association has been seen between level of education attainment and women early marriage. Women who attained primary and higher education level had reduced the probability of early marriage by 13% (OR=0.87, CI: 0.784, 0.966) and 21.1% (OR=0.789, CI: 0.633, 0.983) respectively as compared to non educated women by controlling the effect of random effect and other variables in the model.

Another significant ingredient of women early marriage is wealth index. Women who came from family of richer and middle economic status were 21.7% (OR=1.217, CI: 1.002, 1.477) and 24.5% (OR=1.245, CI: 1.02, 1.518) more likely to be early married than those came from the richest family respectively by controlling the effects of random effect and other covariate in the model. Similarly the odds of women early marriage those came from poorer family were 1.325 (OR=1.325, CI: 1.086, 1.616) times more likely to be early married as compared to those from the richest family. In Ethiopia husband educational attainment have statistical significant association with women early marriage. Thus, the result of this study revealed that the odds ratio of women being early married were reduced by 15.2% (OR=0.848, CI: 0.732, 0.981) and 41.8% (OR=0.582, CI: 0.491, 0.689) for women having husband secondary and higher education level as compared to women having non-educated husband respectively by keeping the random effect and all other covariate constant.

Husband occupation status also another influential predictor variable, for women early marriage. The odds ratio of women being early married were reduced by 20.2% (OR=0.798, CI: 0.691, 0.923) and 14.7% (OR=0.853, CI: 0.756, 0.963) for women having husband business worker and other work category (i.e. excluding agriculture, professional, business and laborers) respectively as compared to women having farmer husband keeping the random effect and other variables in the model constant. The total number of sibling where women came from is another significant factor of women early marriage in Ethiopia. The chance of women early marriage increase as the number of family size increased. Thus, the odds ratio of women early marriage was increased by 10.7% (OR=1.107, CI: 1.016,1.205) for women came from total number of sibling greater than five as compared to less family size.

The Predicted Probability of Women Early Marriage

The fitted line for a given regions would differ from the average line in its intercept, by an amount of random effect U_{0j} for region j. A plot of the predicted region lines would, therefore, show a set of parallel lines. To produce this plot, we use the predicted log odds of women early marriage for each women, based on their place of residence, religion of respondent, women education level, wealth index, husband education level, husband occupation status and total number of sibling with their respective region of residence.

This is given by $\frac{\exp(\hat{\pi})}{1+\exp(\hat{\pi})}$, where, $\hat{\pi}$ is the predicted log odds of women early marriage in each predictor variable. Therefore, variables, having high regional effect, would be considered as regionally varied indicators of women early marriage. Accordingly, place of residence and women education level were identified having high regional effect and further considered as regionally varied indicators of women early marriage shown on figure 4.4 in Appendix.

Place of Residence

For married women with their place of residence, the predicted log-odds of being early married ranges from about -0.9 to 1.4 depending on the region of residence. This translates to a range in probabilities of $\frac{\exp(-0.9)}{1+\exp(-0.9)}$ =0.28 to $\frac{\exp(1.4)}{1+\exp(1.4)}$ = 0.802. So there are strong regional effects between place of residence (Appendix: Fig 4.4). Place of residence is supposed to be regionally varied variables have high random effects on women early marriage across regions.

Women Education level

For married women with their education level, the predicted log-odds of being married early ranges from about -1.25 to 1.3 depending on the region of residence. This translates to a range in probabilities of $\frac{\exp(-1.25)}{1+\exp(-1.25)}$ =0.222 to $\frac{\exp(1.3)}{1+\exp(1.3)}$ = 0.785. So there are strong regional effects between women education level (Appendix: Fig 4.5). Thus women education level is supposed to be regionally varied variables have high random effects on women early marriage across regions.

4.3.4 The Results for Random Intercept with Random Coefficient Model

So far, we have allowed the probability of women early marriage vary across regions, assuming that the effects of the explanatory variables are the same for each region. But, it is essential to determine whether the explanatory variables included in the study have different influence on the response variable (early married women) by varying them among regions. Now we are going to see the effect of each associated factors based on predicted probability of women early marriage versus regional effects. As a result place of residence and women education level have high effects on predicting women early marriage and supposed to be vary across the region. From the results of random slope multilevel model analysis the fixed part of the model clearly show the relationship between women early marriage and its associated factors. Thus, place of residence, religion of respondent, women education attainment, wealth index, husband's educational status, husband's occupation status and total number of sibling were found to be statistically significant determinants of women early marriage in Ethiopia**.**

It also shows that, in the random effect part, the value of 0.0448 is the estimated variance of intercept (region). The log odd of women early marriage in an average region with $(U_j = 0)$ is estimated as $\hat{\beta}_0 = 0.291$ that is shared by all regions. The log-odds of the probability of women early marriage for region j is given by $0.291 + \widehat{U}_j$ where the variance of the intercepts across region is estimated as $\mathrm{var}(\mathrm{U_j}) = 0.0448$, which is referred to as between region variance. In order to know the statistical significance of supposed random variables included in the model the random slope test takes place.

The random effect tests examine hypothesis that whether or not the random slope or the variance of included random coefficient is needed for this model to fit the data in order to show variation of women early marriage across region.

Hypothesis test for random effect part:

$$\mathrm{H_0}\!:\ \omega = \begin{bmatrix} \sigma_{u0}^{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ vs } \mathrm{H_A}\!:\ \omega \text{ is a (3X3) positive semi definite.}$$

Because the constrained variance component test lies on the boundary of the parameter space, the likelihood ratio test can break down asymptotically. It has been shown that tests for variance component can be carried out using mixtures of chi-square distributions. In this study, we show that the null distribution of this one sided likelihood ratio test statistic converges to a 50:50 mixture of chi-square distributions with 1 and 5 degree of freedom given as:

$$0.5X_1^{\,2}+0.5X_5^{\,2}$$

$$\mathrm{p-value} \ = 0.5\Pr(\mathrm{X^2}_{1:5} > Likeli\mathrm{hood\ ratio\ test})$$

$$= 0.5\Pr(\mathrm{X^2}_1 > 5.7434) + 0.5\Pr(\mathrm{X^2}_5 > 5.7434) = 0.1742$$

The likelihood ratio test statistic is equal to twice the difference of the log likelihoods of random intercept model with fixed coefficient and the log likelihoods of random slope model, or the difference of likelihood under null hypothesis and alternative hypothesis respectively which stated as:

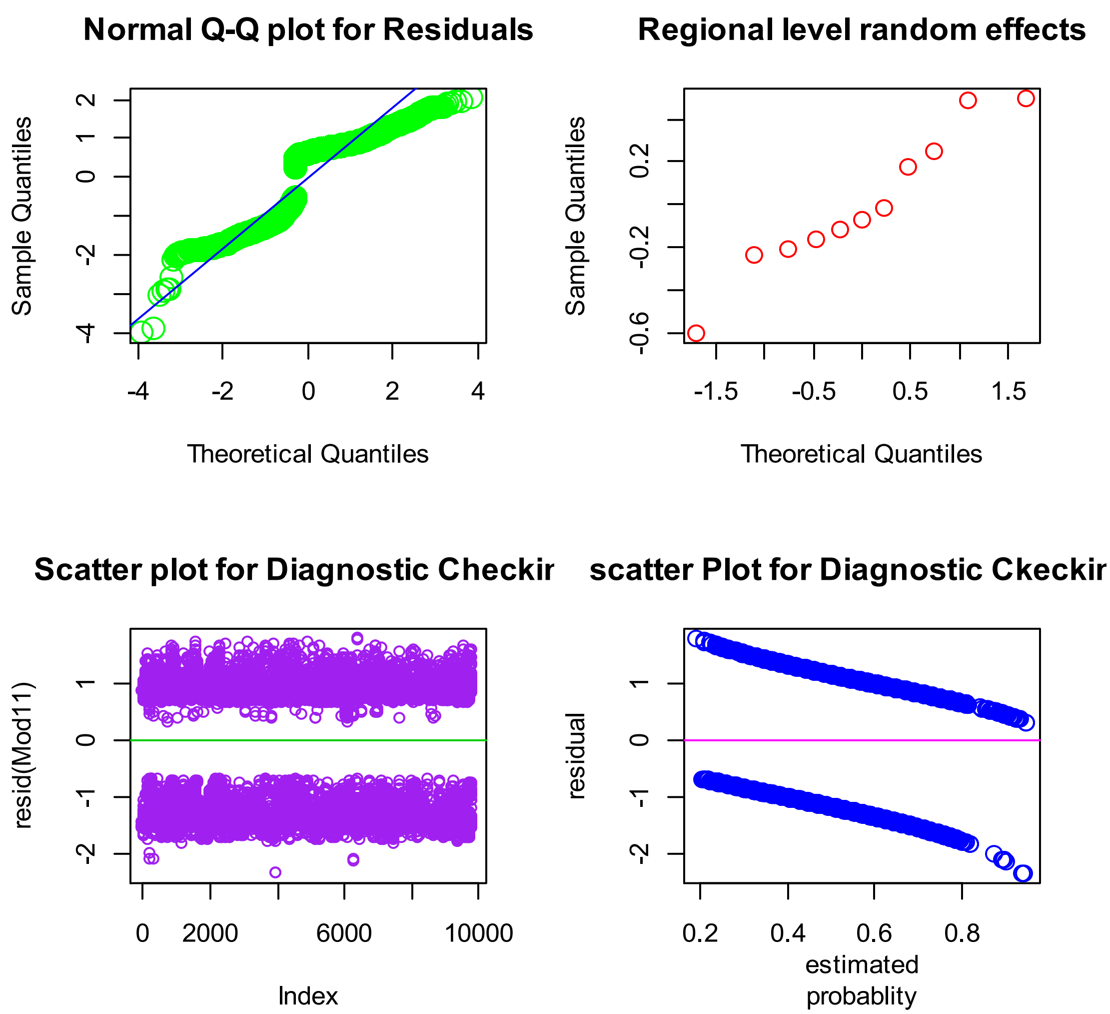
$$\mathrm{LRT} = -2\{l(Y|H_0) - l(Y|H_1)\}$$

Where $l(Y|H_0)$ and $l(Y|H_1)$ are the log likelihoods under the null and alternative hypotheses evaluated at their maximum likelihood estimates, respectively. The critical value for $\alpha = 0.05$ test using this mixture of chi-square distribution with 1 and 5 degree of freedom is 7.455, indicating that we retain H_0 for LRT < 7.455 and p-value = 0.1742 > 0.05. Estimates of this model show that the estimated variances of random slopes of all included variables are zero. This indicated that adding the random explanatory variables to the model was not necessary in order to detect variation of women early marriage among regions. The results confirms that the variation of women early marriage due to the effect of place of residence and women education level do not significantly different from zero across region.

4.3.5 Multilevel Model Diagnosis

Model diagnostics are used to detect problems with the model and suggest improvements. A failure to detect outliers and influential cases can have severe distortion on the validity of the inferences. The diagnostic plot for residuals like the normality for Pearson and standardized residuals of the multilevel model were presented under Figure 4.4. Therefore Q-Q plot for normality of residual and scatter plots of diagnostic checking model was performed. The Q-Q plot from the following Figure in first panel verifies that the residuals are close to normally distributed. Q-Q plots for normality of random effects at regional levels are also given in the figure at panel two, and illustrates that the intercept (regional) random effects are normally distributed with mean zero and variance σ^2_{u0} . And also scatter plot from the Figure in third and forth panel verifies that the residual of the model versus index or residual versus the estimated probability did not show any systematic pattern. This suggested that the residuals are symmetric around zero (i.e. positive and negative residuals are almost equal). For acceptable fit one would expect that locally the residual average zero, the smooth line helps in detecting a deviation from this expectation. Therefore, from these residual plots the model fit is very well. Thus, the fitted multilevel logistic regression model is good fit for the given data.

Figure 4.2 Diagnosis plots for the Multilevel Logistic Regression Model



4.4 DISCUSSIONS

This study attempted to determine the socio-economic and demographic associated factors of early marriage among women in Ethiopian. The study also aimed to investigate the regional variability of women early marriage by identifying factors that accountable for variation across regions of Ethiopia using data from EDHS, 2016.

The detail discussion of the results was based on the output obtained from descriptive, single and multilevel model analysis. The results of the study showed that, out of a sample of 9825 married women from 11 regions considered, 5976 (60.8%) women were married at early age while 3849 (39.2%) were married at the age of 18 years and above.

Based on the results of this study, women who lived in Amhara and Affar were more likely exposed to early marriage compared to women who lived in Tigray where as women who lived in Addis Ababa, Dire Dawa, Somale and SNNP were less likely exposed to early marriage than those lived in Tigray region. This could be due to the desire of the family to keep one's good name and social esteem and (mostly of fathers), to be seen as a means of ensuring her and families safety. The other possible reason might be Tradition and cultural values of the regions where delayed marriage would not be acceptable in the eyes of the community. Therefore region has statistically significant effect on early marriage. This finding is in agreement with the study conducted in Ethiopia (Assefa *et al.,* 2005; Sileshi *et al.,* 2015) which revealed that the prevalence rates of women early marriage vary greatly by region and are often higher in Amhara region, Ethiopia, where almost 50 percent of girls are married by age 15.

Multilevel logistic regression models were employed to analyze factors that affect women early marriage. Multilevel logistic regression model allows for identifying variation of women early marriage among region while logistic regression model limit to detect the heterogeneity of early marriage between regional-level. Therefore before the analysis of data using multilevel approach, heterogeneity status of women early marriage with regard to regions was checked using chi-square test and it was statistically significant.

The overall variance resulted from empty multilevel model was statistically significant suggesting that the proportion of women early marriage varied across region of Ethiopia. In addition to empty model, the overall variance of constant term in random intercept with fixed coefficient and the random slope multilevel logistic regression model were found to be statistically significant implying that the proportion of women early marriage differs across region. The result was in line with findings of (Assefa *et al.,* 2005; Sileshi *et al.,* 2015). Assefa *et al.,* (2005) suggested that women with the same characteristics in two different regions have different age at first marriage which might be because of the fact that differences in lifestyle, culture, ethnic or environmental determinants between different regions that confirm early marriage differs significantly by region of residence.

In order to make the model comparison, the researcher preferred the likelihood ratio test and the information criteria (AIC and BIC) technique because likelihood ratio test is the appropriate method for hierarchical model comparison that analysis the nested type data. The model comparison takes place among the three candidate multilevel logistic regression models to get the best fitted model. Consequently the results of model comparison realize that random intercept with fixed coefficient logistic regression model found to be the best fitted model to predict the associated factors of women early marriage in Ethiopia. Analysis of the final model indicated that, the effects of place of residence, religion, educational level of women; wealth index, husband's education level, husband’s occupation status and total number of sibling were found to be statistically significant determinants of women early marriage in Ethiopia at 5% level of significance. This finding was agreed with the findings of Mohammed (2018).

Place of residence is a significant factor determining women early marriage in Ethiopia. Women who resided in the rural areas were more likely to occur early married as compared with those from the urban areas (OR=1.195). This finding seems to be consistent with other studies (Annabel Erulkar, 2013; Sileshi *et al.,* 2015). Using logistic regression model Annabel Erulkar (2013) found that among women married before age 15, 82% resided in rural areas of Ethiopia and 79% had never been to school. Similarly Adebowale *et al.,* (2012) used Chi-square and Cox proportional hazard models and showed that in Nigeria women who reside in rural area (H.R=1.15) married earlier than their counterpart in urban area. This could be rural areas tend to have institutional and normative structures such as the kinship and extended family that promote early marriage and it might be the fact that women in urban areas might highly participate or attain education when compared with women in rural areas that resulted them to develop skills, gain resources and achieve maturity to manage an independent household and thus they might be delay marriage.

The finding also revealed that the prevalence of women early marriage were high in women those believes in Muslims and others religion as compared to orthodox followers. This finding is consistent with Sarker (2010) and Rodgers B. (2012). By using logistic regression model Sarker (2010) found that, in Bangladesh there were 92% of Muslim women in the country married earlier compared with 84% of Christianity and 85% of Buddhism. This is because child marriage is rooted in religious and cultural traditions based around protecting a girl’s honor where relation before marriage seen as an extremely ashamed. Using similar model Belinda (2015) also confirmed that among 12 Sub-Saharan African countries certain religious affiliations were positively associated with child marriage where its prevalence was higher among women who believed in Islamic religion, traditional religions or no religion than those women who were Christians.

Women who had primary and higher education attainment were less likely to be early married than women with no educational. These results provide empirical evidence that a woman's educational level is an important determinant of early marriage in Ethiopia. Less or no education leads to increased early marriage and therefore lower levels of education are associated with a higher probability of early marriage and this shows education attainment and prevalence of women early marriage was inversely related. Similar finding by Peninah *et al.,* (2011) in western Uganda. using Cox's proportional hazard model they found that the risk of early marriage for women with primary and secondary education were 18% and 34% lower as compared to non-educated women respectively and conclude that education has a statistically significant and strong delaying effect on age at first marriage. A lower risk of getting early married among educated women may be due to waiting time for schooling and understanding the side effect of early marriage. It is also in line with the study by CSA (2012) found in Ethiopia age at first marriage greatly increases with education; women with more than secondary education get married almost eight years later than those with no education.

Similarly in Ethiopia the association of women early marriage and husband education attainment was statistically significant. The likelihood of women early marriage was decreased as the level of educational attainment of their husband increased. Though the study showed that women whose their husband had secondary and higher education attainment were less likely to be early married than those women with non educated husband. The same finding was done in Bangladesh (Zahangir and Kamal, 2011). Using multiple logistic regression model they found, no/less educated women, marry with no/less educated husbands and it is argued that higher educational attainment was the main force underlying the delay age at first marriage among females.

The finding of the study showed that the risk of women early marriage were increased for women came from the economic status of poorer, medium and richer family as compared to those came from the richest parent. This finding is consistent with the study conducted by (Sileshi *et al.,* 2015). This is because child marriage is most common among the poor who have fewer resources and opportunities to invest in alternative options for girls. And also when parents marry off their daughter, there are often economic and social reasons for them to make that choice. They state that in Ethiopia marriage taken as the way to improve family's economic status. The finding also in line with the study conducted by Hotchkiss *et al.,* (2016) in Serbia and shows that about 24.3% of females living in the poorest quintile of HHs were married by age 15, compared to 12.4% of those in the middle wealth group and 3% of those in the richest wealth groups.

The finding of this study shows that a woman who's their husbands were business worker and other work (i.e. excluding agriculture, professional, business and laborers) have low risk of being early married than that of having farmer husband. The type of work on which the women's husband engaged at the time of their marriage were rooted as the base factor for women early marriage in Ethiopia. This study was in line with the study conducted by (Peninah *et al.,* 2011; Mohammed, 2018). Similarly the probability of women early marriage increase as the number of family size increased. Thus, result revealed that the women who came from the number of sibling greater than five were more exposed than those came from less family size. This finding is consistent with Mohammed (2018) that was family size is significantly associated with women early marriage.

In this study the adequacy of model were checked by using diagnostic plot for residuals like standardized residuals of multilevel model. The residual versus fitted model value plot for final multilevel model presented under multilevel model diagnostics. As a result the plots do not show any systematic patterns that give an idea about the distortion of the model. This points out that the model fits the data well. Thus, we can say that the model was not distorted.

5. CONLUSIONS AND RECOMMENDATIONS

5.1 CONCLUSIONS

In order to identify the socio-economic, demographic and proximate variables that determines women early marriage in Ethiopia, three different multilevel models namely: multilevel with random intercept only, multilevel with random intercept and fixed coefficient and multilevel with random coefficient model have been fitted. Among these random intercept with fixed slope multilevel logistic regression model best fitted the EDHS-2016 data set to predict women early marriage. As a result place of residence, religion, women education attainment, wealth index, husband education attainment; husband occupation status and total number of sibling were identified as the most determinant factors of women early marriage in Ethiopia. Although, there were high variability in proportion of women early marriage between regions of Ethiopia. It is observed that women who resides in rural area were (OR = 1.239) times more likely to be early married than those lived in urban. The reason for this disparity among urban- rural residences might be access to education in urban areas compared to rural areas. Generally, the probability of women being early married were high for those residing in rural area, no or low education attainment, living in poor family and unlimited family sizes.

The result of the study also indicated that there was significant variation of women early marriage across 11 regions. The measure of the correlation between two individuals who were in the same region (ICC) identifies that variation in experiencing women early marriage within region was higher than that of between regions. In addition to that, the results of non-parametric approach based on the chi-square test and the parametric approach based on the multilevel logistic regression model without explanatory variables suggest that the proportion of women early marriage were significantly vary among the regions.

Similarly the results of the study show that the effects of all determinant factors of women early marriage (place of residence, religion, women education attainment, wealth index, husband education, husband occupation status and total number of sibling) were uniform throughout the region. In other words, the relationship between early marriage and associated factors included in the study were uniform throughout the regions. Therefore it is possible to conclude that, the variables included in the study are not responsible for explaining variation in women early marriage between regions of Ethiopia.

5.2 RECOMMENDATIONS

As a result of current study, the concerted action should implement to reduce early marriage with strong commitment. Accordingly, all stakeholders at various levels: national, regional, community, family and individual should play a great role. Based on the results of the study the researcher recommends the following points to the government as well as the researchers.

- This study has only investigated the overall variation of women early marriage across regions. But which region is highly practicing women early marriage is not identified in the model. Hence, future study should apply multilevel spatial model to identify the hotspot areas.
- Awareness has to be given for the society on age at the marriage. The education sector can play an effective role in this regard and the awareness need to follow the ordinance of the legal age of marriage because it is the most determinants of health for women and child borne. Moreover, it is advisable to target young women, particularly those with no or little education including primary school girls, with information on reproductive health and to provide them to avoid ultimately early age marriage.
- Religious leaders, key informants and other stake-holders should spread the understanding that religion does not demand women early marriage.
- All regional takes remedial measures on public awareness, design strategy in order to reduce the harmful practice and improve attitudes of stakeholder living in their region towards associated factors of women early marriage.
- Further studies should be conducted in order to identify the most influential determinant factor of women early marriage by including all associated factors of early marriage.

REFERENCE

Assefa, D., Wassie, E., Getahun, M., Berhaneselassie, M. and Melaku, A. (2005) Harmful Traditional Practices for the Ethiopian Health Center Team.

Annabel Erulkar. (March 2013) Early Marriage, Marital Relations and Intimate Partner Violence in Ethiopia.

Adebowale, S.A., Fagbamigbe, F.A., Okareh, T.o. and Lawal, G.o. (2012). Survival analysis of timing of first marriage among women of reproductive age in Nigeria: *regional differences. Afr J Reprod Health* **16 (**4): 95-107.

Alemayehu Bishaw. (2014). The Impact of Education on Rural Women's Participation in Political and Economic Activities. *International Journal of Educational Administration and Policy Studies,* **6**(2): 223-231.

Agresti, A. (1990). Categorical Data Analysis. John Wiley and Sons, New York.

Agresti, A. (2008); An Introduction to Categorical Data Analysis. 2ndEd, Wiley Inc., New York.

Agresti, A. (2002). Categorical Data Analysis. Wiley Inter science, New York

Akaike, H. (1974). A new look at the statistical model identification. IEEE Trans Automatic Control.

Amin, S., Chong, E. and Haberland, N. (2008). Programs to address child marriage: Framing the Problem. *International Journal of Public Health,* **3**(1): 79-82.

Bedasa Tessema, Salie Ayalew and Kasim Mohammed. (2015). Modeling the Determinants of Time-to-age at First Marriage in Ethiopian Women: A Comparison of Various Parametric Shared Frailty Models. *Science Journal of Public Health*, **3**(5): 707-718.

Belinda. (2015). Minimum Marriage Age Laws and the Prevalence of Child Marriage and Adolescent Birth: *Evidence from Sub-Saharan Africa International Perspectives on Sexual and Reproductive Health,* **41**(2):58–68,

Bewick, V., Cheek and Ball. (2005). Statistics review 14: Logistic regression. Critical care.

Boyden, J., Pankhurst., A. and Yisak Tefera. (2013). Harmful Traditional Practices and Child Protection: Contested Understandings and Practices of Female Child Marriage and Circumcision in Ethiopia. Oxford University, Oxford, England.

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models, J. *Am. Statist. Ass.*, **88**: 9-25.

Choe, M.K., S.Thapa and Mishra, V. (2005). Early Marriage and Early motherhood in Nepal. *Journal of Biosocial science* **37**(2):143-62.

Cook, D. (1977). Detection of influential observations in linear regression. Technometric, 19, 15-18

CSA (Central Statistical Authority). (2012). Ethiopia 2011 Demographic and Heath Survey. Federal Democratic Republic of Ethiopia, Addis Ababa, Ethiopia.

CSA (Central statistical Authority). (2016) Federal Democratic Republic of Ethiopia Demographic and health survey, Addis Ababa, Ethiopia

Eshetu Gurmu and Dula Etana. (2012). The Impact of Mass Media on Women's Reproductive Health Behaviour in Ethiopia. Addis Ababa University. Addis Ababa, Ethiopia.

Fitch CA, Kennedy, S., Oakes, JM. and Ruggles, S. (2011). Marriage Formation and Local Economic Opportunity in the United States: A Comparison of Public-use and Restricted Access Census Data Minnesota Population Center.

FDRE (Federal Democratic Republic of Ethiopia). (2012). Assessing Progress towards the Millennium Development Goals: Ethiopia MDGs Report. Addis Ababa, Ethiopia.

Gage, AJ. (2007). Coverage and Effects of Child Marriage Prevention Activities in Amhara Region, Ethiopia.

Gelman, A. (2006). Data Analysis Using Regression and Multilevel/ Hierarchical Models. Columbia University: Cambridge University Press.

Goldstein, H. (2003). Multilevel Statistical Models. 3rdedition London: Arnold; New York: Oxford University Press Inc

Goldstein, H. and Rasbash, J. (1996). Improved approximations for multilevel models with binary responses. *J. Roy. Statist. Soc., Series* B **159**:505–513.

Goldstein, H. (1991). Nonlinear multilevel models with an application to discrete response data. *J. Biometrika*. **78**: 45-51.

Gordon, B. (2012). Combating child marriage through education. The Office of Gordon and Sarah Brown, London, UK.

G. Rodr´ıguez, (2007). Logit models for binary data.

Gelman, A. and Hill, J. (2006). Data Analysis Using Regression and Multilevel or Hierarchical Models. Columbia University, *Journal of the American Statistical Association*.

Hosmer, D.W. and Lemeshow, S. (2000). Applied Logistic Regression.2nd Ed, John Wiley and Sons Inc., New York.

Hosmer, D. and Lemeshow, S. (2011). Applied Logistic Regression (5th Edition). New York: John Wiley and Sons.

Hotchkiss, D., Godha, D., Gage, A. and Cappa, C. (2016). Risk Factors Associated With The Practice Of Child Marriage Among Roma Girls in Serbia. *Journal of International Health and Human Rights,* **4**(3): 102-109.

Hox, J. (2002). Multilevel Analysis Techniques and Application.

Hox, J. (2010). Multilevel Analysis Techniques and Applications.2nd Ed. New York. Utrecht University. Routledge.

Ikamari, L. (2005). The effect of education on the timing of marriage in Kenya, *Demographic Research,* **12** (1):1-27.

ICF International, C. (2011). Ethiopia Demographic and Health Survey, Central Statistical Agency Addis Ababa, Ethiopia.

ICRW. (2017). Child Marriage in Ethiopia, Nepal, and Niger: Insights from Qualitative Data. Washington, DC: International Center for Research on Women. [Tentative title].

ICRW(International Center for Research on Women). (2015). Child Marriage around the World. ICRW, Washington DC, USA.

ICRW (International Center for Research on Women). 2010. Too young to Wed: Education and Action towards Ending Child Marriage. ICRW, Washington DC, USA.

Joseph N., M.Fajar R. and Mayang R. (2012). Prevalence of Child Marriage and Its Determinants among Young women in Indonesia, Child poverty and social protection Conference. *Journal of Statistics* **14**: 19 - 25.

Kanta, B. (2013). Child marriage in Ethiopia- A matter of concern. *International journal of Management and Social Sciences*, **2**(5): 62-74.

Khan, H. and Shaw, E. (2011). Multilevel logistic regression analysis applied to binary contraceptive prevalence data, University of Warwick.

Kyari, G. V. and Ayodele, J. 2014. The Socio-Economic Effect of Early Marriage in North Western Nigeria. *Mediterranean Journal of Social Sciences*, **5**(1): 116-123.

Longford, N.T. (1993). Random Coefficient Models. Oxford, Clarendom Press.

Mengistu, M. M. (2017). Early Marriage in Ethiopia: So Little Done But So Much to Do. Business and Management, 9(3).

Menard Scott. (2002). Applied Logistic Regression Analysis. 2nd edition. Sage University Paper on Quantitative Applications in the Social Sciencies.No.07-106. Beverly Halls, CA: Sage.

Mohammed (2018) Spatiotemporal Variation of Early Marriage in Babile Woreda, Eastern Hararghe Zone, Oromiya National Regional State, Ethiopia

Myers, J. and Harvey, R. (2011). Breaking Vows: Early and Forced Marriage and Girls' Education.

NCTPE (National Committee on Traditional Practices of Ethiopia). (2003). Ethiopia harmful traditional practices: Old beyond imaginings Addis Ababa.

Nasrin, S. O. and Rahman, K. M. (2012). Factors affecting early marriage and early conception of women: A case of slum areas in Rajshahi City, Bangladesh. *International Journal of Sociology and Anthropology,* **4**(2), 54-62.

Nanda, P., P. Das., N. Datta. and S. Lamba, E. (2015). Making Change with Cash? Evaluation of a Conditional Cash Transfer Program to Improve the Status of Girls in Northern India, Research Brief, Washington, DC: International Center for Research on Women.

Parsons, J., Edmeades, J., Kes, A., Petroni, S., Sexton, M. and Wodon, Q. (2015). Economic impacts of child marriage: a review of the literature. *The Review of Faith & International Affairs,* **13**(3), 12-22.

Peninah, A., Leonard, K., Atuhaire and Gideon Rutaremwa. (2011). Determinants of age at first marriage among women in western Uganda. Presented in European Population Conference 2010, Vienna 1-4 September 2010

Raudenbush, S. and Bryk, A.S. (1992). Hierarchical Linear Models. Newbury Parks, CA: Sage.

Raudenbush, S.W., Yang, M. L. and Yosef, M. (2000). Maximum likelihood for generalized linear models with nested random effects via high-order, multivariate Laplace approximation, *Journal of Computational and Graphical Statistics* **9**: 141 – 157.

Robert and Rebecca. (2010). Early Female Marriage in the Developing World

Sarker, P. (2010). Determinants and effects of Early Marriage in Bangladesh. *Research Journal of Applied Science*, **4**(5): 73-79.

Sinha, K. (2009). Almost half of Indian women married as minors: Study. The Times of India. Bennett, Coleman & Co. Ltd.

Shiferaw, S., Spigt, M., Godefrooij, M., Melkamu, Y. and Tekie, M. (2013). Why do women prefer home births in Ethiopia? BMC pregnancy and childbirth, 13(1), 5.

Sofia, N. and Khalid, M., (2015) Causes and Consequences of Child Marriages in South Asia*. A Research Journal of South Asian Studies,* **30**(2): 961 – 975.

Sileshi Workineh, Getiye Dajane and Genet Dagu. (2015). Determinants of Early Marriage Among Female Children in Sinan Woreda North West Ethiopia. *Health Science Journal,* **9**(6): 4-5.

Snijders, T. and Bosker, R. (1999). Multilevel Analysis: an Introduction to Basic and Advanced Multilevel Modeling. London/ Thousand Oaks/ New Delhi: Sage Publications.

UNIFPA. (2006). In ending child marriage, A guide for global policy action International Planned Parenthood Federation and the Forum on Marriage and the Rights of Women and Girls.

UN. (2007). The Millennium Development Goals, Report 2007, United Nations, New York. Retrieved 12, August, 2013 http://www.un.org/mi1lenniumgoals/pdf/mdg2007.pdf

UNICEF (United Nations Children's Fund). (2014). Ending Child Marriage: Progress and Prospects. New York: UNICEF

UN (United Nations). (2008). Women Watch: Ethiopia National Action Plan. United Nations, New York, USA.

WHO. (2006). Constitution of the World Health Organization. www.who.int/governance/eb/ who_constitution_en.pdf.

Wodon, Q. T., C. Male, K. A. Nayihouba, A. O. Onagoruwa, A. Savadogo, A. Yedan, J. Edmeades, A. Kes, N. John, L. Murithi, M. Steinhaus and S. Petroni. (2017). Economic Impacts of Child Marriage: Global Synthesis Report. Economic Impacts of Child Marriage. Washington, D.C.: World Bank Group.

Zahangir, M.S. and Kamal, M.M. (2011**)**. Several Attributes Linked with Child Marriage of Females' in Bangladesh, *International Journal of Statistics and Systems,* **volume 6**.

Zoran Bursac, C. Heath Gauss, D. Keith Williams and David Hosmer. (2007). A Purposeful Selection of variables macro for Logistic Regression, University of Arkansas for Medical Sciences and University of Massachusetts. Paper 173.

APPENDIX

Table 4.7 Test of Significance Level of Individual Predictors for Binary Logistic Regression Using Score Test

Variables	Wald chi-square	d.f	P-value
Region	182.004	10	0.001
Place of Residence	5.071	1	0.024
Religion	27.610	3	0.001
Women Education	8.820	3	0.032
Wealth index	10.513	4	0.033
Respondent work	4.174	1	0.176
Media Exposure	2.541	1	0.111
Husband Education	51.242	3	0.001
Husband occupation	12.518	4	0.014
Total No. Sibling	5.677	1	0.017

Table 4.8 Results of maximum likelihood estimates of parameters in fitting binary logistic regression model

Variable	Categories	$\hat{\beta}$	S.E	p-value	OR	[95% C. I]
Intercept		0.2969	0.070	2.6e-5 *	1.636	1.268, 2.112
Region	Tigray(ref.)					
	Affar	0.332	0.124	0.008 **	1.394	1.091, 1.780
	Amhara	0.2046	0.059	0.000 **	1.409	1.161, 1.710
	Oromia	-0.159	0.061	0.009 *	0.770	0.633, 0.938
	Sumale	-0.263	0.071	0.000 ***	0.652	0.519, 0.818
	B/Gumes	-0.121	0.066	0.067	0.819	0.662, 1.012
	SNNPR	-0.243	0.065	0.000***	0.670	0.544, 0.827
	Gambella	0.051	0.073	0.482	1.085	0.857, 1.372
	Harari	-0.198	0.075	0.008 *	0.723	0.568, 0.920
	A/Ababa	-0.523	0.073	<0.001***	0.429	0.339, 0.527
	Dire Dawa	-0.229	0.075	0.002 **	0.689	0.542, 0.875
Place of Residence	Urban(ref.)					
	Rural	0.127	0.055	0.022*	1.225	1.026, 1.462
Religion of	Orthodox(ref.)					
	Muslim	0.103	0.040	0.010*	1.185	1.041, 1.350

Women	Protestant Others	0.022 0.971	0.048 0.190	0.639 <0.001 **	1.041 5.444	0.892, 1.215 2.682, 11.051
Women Education	None (ref.) Primary Secondary Higher	 -0.085 -0.044 -0.142	 0.032 0.053 0.069	 0.009 * 0.413 0.038*	 0.871 0.930 0.790	 0.784, 0.966 0.783, 1.104 0.634, 0.985
Wealth Index	Richest(ref.) Richer Middle Poorer Poorest	 0.115 0.130 0.169 0.065	 0.060 0.062 0.062 0.060	 0.057 0.036* 0.006** 0.282	 1.209 1.238 1.320 1.111	 0.996, 1.468 1.015, 1.511 1.082, 1.610 0.916, 1.346
Women work	Working (ref.) Not working	 -0.010	 0.029	 0.708	 0.980	 0.892, 1.076
Media Exposure	Yes(ref.) No	 -0.059	 0.038	 0.118	 0.905	 0.801, 1.023
Husband education level	None (ref.) Primary Secondary Higher	 0.039 -0.098 -0.335	 0.032 0.046 0.053	 0.213 0.032* <0.001 *	 1.065 0.852 0.584	 0.962, 1.186 0.736, 0.986 0.493, 0.691
Husband Occupation	Agriculture(ref.) Professional Business Labourers Others	 -0.038 -0.136 -0.185 -0.096	 0.038 0.045 0.158 0.038	 0.308 0.002 ** 0.240 0.011 **	 0.935 0.800 0.740 0.854	 0.828, 1.057 0.692, 0.925 0.448, 1.221 0.756, 0.964
No. of sibling	< = 5 (ref.) >5	 0.064	 0.026	 0.015*	 1.109	 1.018, 1.208

Note _*'indicates significance for $p<0.05$: The reference categories were selected subjectively considering previous research.

Table 4.9 Output of Random slope multilevel Logistic Regression Analysis

Covariates	Categories	Estimate	Std. Err.	P-value	OR	[95%C.Ifor OR]
Intercept	Constant	0.291	0.119	0.015*	1.337	1.058, 1.690
Place of Residence	Urban(ref.)					
	Rural	0.249	0.100	0.013*	1.284	1.054, 1.563
Religion of Women	Orthodox(ref.)					
	Muslim	0.145	0.065	0.027*	1.157	1.017, 1.316
	Protestant	0.042	0.078	0.593	1.042	0.893, 1.217
	Others	1.679	0.361	0.0001*	5.361	2.640, 10.885
Women Education	None (ref.)					
	Primary	-0.141	0.059	0.017*	0.868	0.773, 0.975
	Secondary	-0.067	0.102	0.511	0.935	0.765, 1.142
	Higher	-0.189	0.140	0.177	0.827	0.628, 1.089
Wealth Index	Richest(ref.)					
	Richer	0.182	0.099	0.066	1.200	0.987, 1.458
	Middle	0.204	0.101	0.045*	1.227	1.004, 1.498
	Poorer	0.265	0.101	0.009*	1.304	1.068, 1.592
	Poorest	0.102	0.099	0.299	1.108	0.912, 1.345
Women work	Working (ref.)					
	Not working	-0.019	0.047	0.686	0 .980	0.892 1.077
Media Exposure	Yes(ref.)					
	No	-0.100	0.062	0.110	0.904	0.800, 1.022
Husband education level	None (ref.)					
	Primary	0.639	0.053	0.231	1.066	0.960, 1.183
	Secondary	-0.155	0.074	0.037*	0 .855	0.739, 0.990
	Higher	-0.546	0.086	0.0001*	0.578	0.488, 0.685
Husband Occupation	Agriculture(ref.)					
	Professional	-0.063	0.062	0.311	0.938	0.830, 1.061
	Business	-0.220	0.074	0.003*	0 .801	0.693, 0.927
	Labourers	0.299	0.255	0.241	0 .740	0.448, 1.222
	Others	0.161	0.062	0.009*	0 .850	0.753, 0.960
No. of sibling	< = 5 (ref.)					
	>5	0.098	0.043	0.024*	1.103	1.013, 1.202

Random effect	Variance component	S.E	95% Confidence Interval	
$\sigma_0{}^2 = \mathrm{var}(U_{0j})$	0.0448	0.0654	0.0025,	0.7809
$\sigma_1{}^2 = \mathrm{var}(U_{1j})$	0.0137	0.0208	0.0007,	0.2683
$\sigma_2{}^2 = \mathrm{var}(U_{2j})$	0 .0070	0.0069	0.0010,	0.0492
$\sigma_{01} = \mathrm{cov}(U_{0j}, U_{1j})$	-0 .0033	0.0331	-0.0683,	0.0616
$\sigma_{02} = \mathrm{cov}(U_{0j}, U_{2j})$	0.0176	0.0127	-0.0073,	0.0426
$\sigma_{12} = \mathrm{cov}(U_{1j}, U_{2j})$	-0 .0026	0.0083	-0.0189,	0.0136

Where $\mathrm{var}(U_{0j})$-Variance of intercept,$\mathrm{var}(U_{1j})$ -variance of place of residence, $\mathrm{var}(U_{2j})$ - variance of women education, $\mathrm{cov}(U_{0j}, U_{1j})$ - covariances of intercept and place of residence and $\mathrm{cov}(U_{0j}, U_{2j})$ - covariance's of region and women education.

Figure 4.3 Predicted Probability of Early Marriage by Place of Residence vs Region

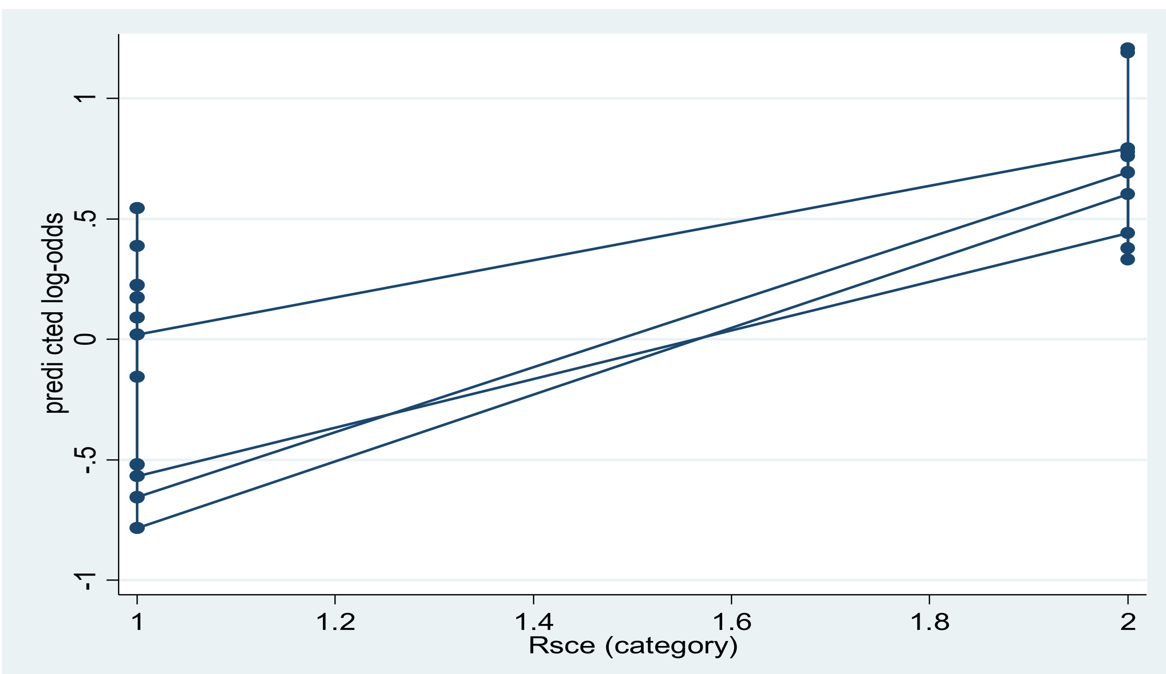


Figure 4.4 predicted probability of early marriage by place of residence vs. region