JIMMA UNIVERSITY
COLLEGE OF NATURAL SCIENCE
SCHOOL OF GRADUATE STUDIES
DEPARTMENT OF STATISTICS

MODELING DELIVERY CARE SERVICE UTILIZATION OF MOTHERS IN ETHIOPIA

By: Amsalu Arega

A Thesis Submitted to Department of Statistics, School of Graduate Studies, College of Natural Science, Jimma University, in Partial Fulfillment for the Requirement of Masters of Science in Biostatistics

January 2014
Jimma, Ethiopia

MODELING DELIVERY CARE SERVICE UTILIZATION OF MOTHERS IN ETHIOPIA

By: Amsalu Arega

Advisor: Legesse Negash (PhD Scholar)

Co-advisor: Ashenafi Abebe (MSc.)

January 2014
Jimma, Ethiopia

STATEMENT OF AUTHOR

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 in the university library to be made available to borrowers for reference. I earnestly 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.

Amsalu Arega

Date: ___________________

Signature: ______________

Jimma University, Jimma

ACKNOWLEDGMENT

First, I wish thanks to GOD as all is and has been possible through him. Secondly, I am very thankful to my main advisor, Legesse Negash (PhD Scholar), who provided valuable guidance and inspiration starting from selection of title up to this moment. Throughout my work, he provided encouragement, sound advice, and good teaching. I would like to thank also my co-advisor Ashenafi Abebe (MSc.) who gives me all the necessary advices and comments regarding my doings.

I would like to pass my heartfelt gratitude and pleasure to my friends Kassu Mehari and Yimer Wasihun who gave me moral appreciation and fruit full guidance throughout my career. I would like also to thank the Ethiopian Central Statistics Agency for providing me with all the relevant secondary data used in this study.

Last but not least, I would like to thank my parents who have invested all their life to hold up me. My parents, I would like to thank you vastly for your sweet words at every time that gave me a long power and anticipate.

DEDICATION

This thesis is dedicated to the memory of my beloved mother, Brkie Manaye; she was passed away due to the complication of delivery.

.

Amsalu Arega

Date: ____________________

Signature: _______________

Jimma University, Jimma

ABSTRACT

Delivery care through access to health facilities and skilled health personnel are the main important interventions for safe motherhood. The presence of a trained health-care worker during delivery is vital in reducing maternal deaths. Delivery assisted by skilled providers is the most important proven intervention in reducing maternal mortality and one of the MDG indicators to track national effort towards safe motherhood. In Ethiopia, the proportion of births attended by a skilled health professional and delivered in a health facility has remained around 6% over the past five years. Increasing the proportion of births delivered in a health facility and under the supervision of health professionals is important to reducing health risks among mothers and children. The main objectives of this study was modeling delivery care service utilization of mothers using marginal and generalized linear mixed models as well as evaluate the determinant factors for the delivery care service utilization of mothers in Ethiopia. Data was taken from the 2011 Ethiopian demographic and health survey, which is a nationally representative survey of mothers in the 15-49 years age groups. Two model families, generalized estimating equation and alternating logistic regression models from marginal model family, and generalized linear mixed model from cluster specific model family were used for the analysis. The result showed that only 17.2% of the mothers received assistance during delivery from health professional. Alternating logistic regression model was best fits the data for population-averaged effects of the given factors on delivery service utilization than generalized estimating equation model and generalized linear mixed model with two intercepts was the best model to evaluate within and between regional heterogeneity of delivery service utilization. All the fitted models gave the same conclusion that age, place of residence, mother's education level, religion, wealth index, birth order, partner's education level and exposure to mass media are the most determinant factors of delivery service utilization of mothers. We conclude that education and wealth quintile have a positive association with delivery assistance where as birth order has a strong negative association with delivery service utilization.

Table of Contents

ACRONYMS

ALR	Alternating logistic regression
AIC	Akaike's information criterion
CSA	Central Statistical Agency
DHS	Demographic and Health Survey
EDHS	Ethiopian Demographic and Health Survey
ESPS	Ethiopian Society of Population Study
GEE	Generalized Estimating Equation
GLM	Generalized Linear Model
GLMM	Generalized Linear Mixed Model
MDG	Millennium development goal
ML	Maximum Likelihood
UNFPA	United Nations Population Fund
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WHO	World Health Organization

LIST OF TABLES

Tables' **page**

CHAPTER ONE
INTRODUCTION

1.1 BACKGROUND

Delivery care through access to health facilities and skilled health personnel are the main important interventions for safe motherhood. Historically, increasing women's access to health facilities with the capacity to provide emergency obstetric care has been responsible for large drops in maternal mortality (Wang *et al*, 2011).

Proper medical attention and hygienic conditions during delivery can reduce the risk of complications and infections that can cause the death or serious illness of the mother and the newborn baby. The presence of a trained health-care worker during delivery is vital in reducing maternal deaths. A skilled health professional can administer interventions to prevent and manage life-threatening complications, such as heavy bleeding, or refer the patient to a higher level of care when needed. The type of assistance a mother receives during childbirth has important health consequences for both mother and child. In addition, the proportion of births attended by skilled providers is a measure of the health system's effectiveness, accessibility, and quality of care. Delivery assisted by skilled providers is the most important proven intervention in reducing maternal mortality and one of the MDG indicators to track national effort towards safe motherhood (Baral *et al*, 2010). An important component of efforts to reduce health risks to mothers and children is increasing the proportion of babies that delivered in health facilities (EDHS, 2011).

According to estimates by WHO, UNICEF, UNFPA, and the World Bank, 358,000 maternal deaths occurred worldwide from preventable complications during pregnancy and childbirth. Moreover, 99% of maternal deaths (355,000) in 2008 occurred in developing countries, and an estimated 87% (313,000) occurred in sub-Saharan Africa and South Asia (WHO, 2008). In developing regions overall, the proportion of deliveries attended by skilled health personnel rose from 55% in 1990 to 65% in 2009. Despite dramatic progress in many regions, coverage remains lower in sub-Saharan Africa and Southern Asia, where the majority of maternal deaths occur. In half of the 38 countries analyzed, the majority of women delivered their last child in an institutional setting. In North Africa/West Asia/Europe and in Latin America and the Caribbean, women are more likely to give birth in a health facility than are women in sub-Saharan Africa or South/Southeast Asia. In about half of the countries, most births take place in a health facility, but the percentage is around 50% in several countries, including Mozambique, Rwanda, Tanzania, and Zambia. The region's countries with the lowest levels of births in health facilities are Ethiopia 6%, Chad 13% and Niger 18%, (Wang *et al*, 2011).

In Ethiopia, the proportion of births that occur at home remains higher and skilled health professionals attended births are very few. The proportion of births attended by a skilled health professional and delivered in a health facility has remained around 6% over the past five years, a far lower level than in other African countries, such as Cameroon 62%, Senegal 62%, Malawi 57%, and Lesotho 52% (MI, 2007).

According to 2011 Ethiopian demographic and health survey (EDHS), only 10% of births have delivered at a health facility (9% in a public facility and 1% in a private facility). Nine women in every ten have delivered at home. The percentage of deliveries in a health facility doubled from 5% the 2005 EDHS, while home deliveries decreased slightly from 94% to the current level of 90%. A skilled provider (4% by a doctor and 7% by a nurse or midwifery) assisted 10% of births, a health educated workers (HEW) assisted less than a relative, or some other person assisted only 1% of births. Similarly, 28% of births have assisted by a traditional birth attendant, while 4% of births were unattended. Skilled assistance at delivery increased from 6 to 10% in the last six years (EDHS, 2011).

The safe motherhood initiative strongly emphasizes ensuring the availability and accessibility of skilled care during pregnancy and childbirth, of which institutional delivery is one element. This would avoid most maternal deaths occurring from preventable obstetric complications. However, as previous studies have clearly demonstrated, the utilization of existing maternal health services is very low in Ethiopia (Eyerusalem, 2010, Asmeret, 2013). Increasing the proportion of births delivered in a health facility under the supervision of health professionals is important to reducing health risks of mothers and children (MI, 2007).

Certain data will not be continuous like binary and count data, (in this case binary data), and the corresponding outcome variables are categorical and count responses. Such outcome variables will not be normally distributed rather distributed as binomial, poisson, gamma etc.

In addition, in case of multistage or clustered sampling procedure, responses variables will be correlated within individuals in the same clusters. EDHS data is a two stage stratified sampling where mothers are the second sampling unit in each clustered within regions. There may be also having regional variations that is; there may be heterogeneity within regions as well between regions in delivery service utilization. To handle such types of data, the most flexible and appropriate models should be applied. This includes generalized linear models (GLM) and its extension, which are capable of analyzing correlated and non-normal data (i.e binary in this case).

Marginal models as generalized estimating equations (GEE) and alternating logistic regression (ALR) models are an extension of GLM by considering dependency in the response variables for clustered data and repeated measurement (Molenberghs & Verbeke, 2005). Cluster specific models like generalized linear mixed models (GLMM) also are a natural outgrowth of both linear mixed models and generalized linear models (Cosmas, 2011). GLMM can be developed for non-normally distributed responses, allow nonlinear links between the mean of the response and the predictors, and can model over dispersion and correlation by incorporating random effects (McCulloch, 1997). This study applied those models to incorporate the nature of the given data.

To summarize, cross-sectional surveys can serve as tools to collect subset of possible covariates, which can be used to establish association with the response variable of interest. This association in turn enables the researcher to establish evidence based in policy planning and resource allocations and in evaluating progress in policy implementations. This can be reached by applying proper statistical methods in measuring the outcome indicators as well as quantifying the impact of determinant factors, interventions coverage and other possible indicators are essential in the success of any policy.

1.2 STATEMENT OF THE PROBLEM

Delivery service by skilled assistance through delivery is a crucial issue in reducing the risk of complications and infections that can cause the death or serious illness of the mother and the newborn baby. Despite the fact that delivery service utilization is essential for further improvement of maternal and newborn, the coverage of delivery service in Ethiopia is still near to the ground, only 10%. Even if there is physical access to institutional delivery services, many mothers may not use them because of different determinant factors at individual, household, and community levels that shape an individual's ability to seek health care. Moreover, previous studies on this area in Ethiopia were considered about modeling only the fixed effects of covariates without including the random effects and with no considering sampling structures of data. Most of the studies previously done are simply using only the ordinary logistic regression model.

Thus, the little magnitude of this service and lack of appropriateness of the model applied for clustered data have generated interest in assessing determinant factors affecting delivery care service utilization of mothers by fitting a statistical model that can explain the data in most meaningful manner.

This study, therefore, has tried to fill the gaps in understanding the status of mothers using health care services for delivery by identifying determinant factors of delivery service in Ethiopia and assessing the performance of different models using clustered data from EDHS 2011 by addressing the following research questions:

- Which fitted model for the delivery care service utilization is statistically plausible?
- Which covariates are the most determinant factors for delivery care service?
- Is there a significant within and between regional heterogeneity in delivery care service utilization of mothers?

1.3 OBJECTIVE OF THE STUDY

1.3.1 General objective

General objective of the study is modeling delivery care service utilization of mothers using marginal and generalized linear mixed models as well as evaluate the determinant factors for the delivery care service utilization of mothers in Ethiopia.

1.3.2. Specific objectives

The specific objectives of the study, which should be accomplished to achieve the general objective stated above, are:

1) To fit models that yield statistically plausible and interpretable estimates of important covariates on delivery care service utilization for the given data.
2) To compare generalized estimating equation & alternating logistic regression models.
3) To assess determinant factors that may affects the delivery care service utilization of mothers.
4) To assess between and within regional heterogeneity of delivery service utilization of mothers.

1.4. SIGNIFICANCE OF THE STUDY

The results of this study will be very useful in the development of an effective delivery care and reducing maternal and infant mortality risk.

Specifically:

- It is expected that this study would increase the understanding of mothers about the importance of delivery care service utilization in Ethiopia.
- The results of the study might be appraising understanding of policymakers by clarifying the main determinant factors that affecting the delivery service utilization of mothers in Ethiopia.
- The results of this study might give information to concerned bodies in setting policies, strategies and further investigation for increasing delivery service utilization.
- The results can provide an important input for any possible intervention in this area for the future.

CHAPTER TWO

LITERATURE REVIEW

2.1 DELIVERY CARE SERVICE UTILIZATION

Delivery care service: is a service provided to mothers during labor, delivery and the early postpartum period by accredited health professionals (a midwifery, doctor or nurse) who has been educated and trained to proficiency in the skills needed to manage complications in mothers and newborns (Adamu, 2011).

Health professionals or skilled attendants are often available at health facility level, although there is historical evidence of well developed home visiting midwives at community level as in Norway, Sweden and also in Holland. For a mother and her newborn, a skilled birth attendant can make a difference between life and death. Not only can they recognize and prevent medical crises, but can identify obstetric complications early and effect immediate referral is a life saving care (Anna, 2007).

Delivery care service can be divided in to two general categories. The first one is basic care, which includes attending normal deliveries, care of the newborn and immediate stabilization of a mother if she has complications before referral. The second one is emergency obstetric care (EMOC) which can further be divided into basic or comprehensive (Eyerusalem, 2010).

Basic or comprehensive emergency obstetric care includes the following six lifesaving procedures: parenteral (intravenous or intramuscular) oxytocics; parenteral antibiotics; parenteral anticonvulsants or sedative; manual removal of placenta; removal of retained products (e.g. with manual vacuum aspiration kit); and assisted vaginal delivery (e.g., using a vacuum extraction or by forceps delivery). A facility offering all six lifesaving procedures is a basic emergency obstetric care facility.

Comprehensive emergency obstetric care includes the provision of eight lifesaving procedures that includes all the six basic emergency obstetric care functions plus, obstetric surgery such as cesarean section and safe blood transfusion. A facility offering all eight lifesaving procedures is a comprehensive emergency obstetric care facility (Susan & Elizabeth, 2010). Maternal health care service utilization is important for the improvement of both maternal and child health.

According to findings from levels and trends in the use of maternal health services in developing countries, except for a few countries (Benin, Namibia, Zimbabwe and Vietnam), the use of skilled care for delivery is considerably lower in sub-Saharan Africa and South/Southeast Asia than in the other regions (Wang et al, 2011).

In a study of six African countries, lower rates of maternal and neonatal mortality in addition, morbidity were shown to have a positive relationship with giving birth in a health facility with the help of trained medical workers (Stephenson *et al*, 2006). Improving maternal and child health requires increasing the percentage of mothers giving birth in health institutions with the assistance of trained personnel, which is the fundamental goal of the safe motherhood and child survival activities (Kesterton *et al*, 2010).

According to an analysis of DHS data from 48 developing countries since 2003, in 23 countries more than half of the births are reported to take place at home (Montagu *et al*, 2011). Findings on delivery practices among women in rural India, Punjab, showed that more respondents reported that home delivery than reported institutional delivery (Garg *et al*, 2010). Another study in a semi urban settlement of Zaria Northern Nigeria revealed that most women (70%) were delivered at home and that a majority of deliveries (78%) were not supervised by skilled attendant (Idris *et al*, 2006).

In Ethiopia as reported in the 2005 EDHS, the majority of births take place at home in poor hygienic situation, while only 6% are in a health facility and are assisted by trained person (CSA and ORC Macro, 2006). Moreover, a study on utilization of maternal health care services in Ethiopia shows that, only 11.6% of the mothers were gave birth with assistance from health professionals, which includes doctors, nurses and midwives (Eyerusalem, 2010). Another finding from the John Snow Inc, L10K baseline survey conducted in 2009 indicated that, although institutional delivery improved over the four years since the 2005 EDHS, it was only 12% in 2009, and few deliveries were assisted by health extension workers , though, the health extension workers had received in-service training (JSI, 2009).

2.2. DETERMINANT FACTORS OF DELIVERY CARE SERVICE

Previous empirical studies have found that the delivery services is related to factors such as age of mothers, mother education level, wealth index, place of residence, religious background, mother's work status, and others. These factors and model families also will be discussed in turn.

Mothers' age: Age may sometimes serve as a demonstration for the mothers' accumulated knowledge of health care services, which may have a positive influence on the use of health services. Possible explanations for higher use of maternal health care services by older mother could include the fact that mother in this cohort are generally more experienced and knowledgeable about healthcare services and their use, which may improve utilization. Older mothers may also be more confident and have higher household decision-making power than that of the younger mothers, particularly adolescents (Reynolds *et al*, 2006), which will improve their likelihood of health service use. Studies in Nigeria (Adamu, 2011) showed that a significant association was found between mother's age and delivery care service use where young adults (20-34) years are more likely to deliver in a health facilities. In contradict, the study from multilevel model in Nigeria (Babalola and Fatusi, 2009) also reveled that no significant relationship between age and use of skilled assistance during delivery. Another study in Ghana (Abor *et al*, 2011) from results of the probit model reveals that, age of mother shows a significantly positive relationship with delivery service utilization. This implies that, older mothers are more likely to utilize services delivering at a medical facility as compared to younger mothers.

Education**:** It is well recognized that mother's education has a positive impact on delivery service utilization. A study in Peru from DHS data (Elo, 1992) using logistic regression model found quantitatively important and statistically significant effect of mother's education on the use of prenatal care and delivery assistance. In another study in Rwanda (Umurungi, 2010) revealed women's education was an important predictive factor for usage of delivery services. Education is a key determinant of health facility utilization for delivery, frequently because education increases the mothers' autonomy, understanding and decision making power within the household. It is also likely that educated mothers will tend to seek out higher quality services. An important finding was that even for wealthier mothers, education made a significant difference in determining where the mothers would deliver. Educated mothers were significantly more likely to deliver at a health facility. The study in Indonesia (Kistiana, 2009) show that mother's educational level generally has a positive and significant effect on the use of hospitals or health clinics at time of delivery. The strong influence of mothers' education on the utilization of delivery care services is consistent with findings from other studies.

Place of residence: Kistiana also introduced that (Kistiana, 2009), women's place of residence is the most important variables influencing maternal health in care utilization in terms of using a health facility as the place of delivery and supervised by health professionals. This can be explained because availability, accessibility, and even affordability of these health facilities are more difficult in rural areas than in urban areas. Mekonen also explained that, the extent of variation in the use of delivery care services by residence is striking. Mothers from urban areas are more likely to receive assistance during delivery than that of mothers from rural areas (Mekonen, 2002). As explained, place of residence came out to be a strong predictor of use of this services in the logistic regression model. One possible reason for this discrepancy is the inequitable distribution of health care services; most of the health care services are concentrated in the urban areas (Eyerusalem, 2010). Another study's by Asmeret using binary logistic regression model shows, mothers in urban areas were more likely than that of mothers in rural areas to utilize delivery services (Asmeret, 2013).

Wealth index: The positive association between wealth index and delivery service has been reported in several studies. Umurungi determined that wealth index was one of the most predictive factors of service utilization. Accordingly, his study, greater household wealth provides resources and may enable mothers to seek care and the ability to buy health (Umurungi, 2010). The costs of seeking care may act as a significant barrier to mothers from poorer households. The same study in Ghana (Charles *et*, 2011), the regression model suggest that as compared to those in the poorest wealth quintile, those in the poorer wealth quintile are more likely to deliver in a health facility, with those in the middle wealth quintile being more likely to use deliver services at a health facility than the reference group. Additionally, those in the richer and richest quintiles seem more likely to see services delivery at a health facility.

Another finding by Ethiopian Society of Population Studies, significant difference was observed in the likelihood of receiving delivery and postnatal services by mother's wealth index. Mothers from the highest wealth index are 3.69 times more likely to receive delivery care services than those in the middle category. The wealth index is useful for ranking the socioeconomic status of households (ESPS, 2008). Asmeret also found that low socioeconomic status of the mother is an important predictor of home delivery. The finding can be possibly explained by the fact that poor mothers are unlikely to afford the cost of transport and other medical costs. In Ethiopia, even though the service in a health post is given free of charge, it incurs costs when complicated delivery is referred to health centers (Asmeret, 2013).

Birth order: An important factor for delivery service utilization in most studies is birth order. In most sub Saharan African countries, mothers with one child (birth order 1) are substantially more likely to use a skilled health provider for their last delivery, than mothers with four or more children. Higher levels of use of delivery service are associated with lower birth order (Wang *et al*, 2011). Ethiopian Society of Population Studies found from logistic regression model explained that, mothers with higher birth orders of four and above are less likely to seek delivery care at a health facility as compared to the second and third birth orders (ESPS, 2008). According to Mesfin *et al* birth order is another obstetric factor found to be significantly affecting the use of safe delivery services (Mesfin *et al*, 2004). The probability of giving birth at health facilities decreased in third and above births.

Exposure to Mass media: Women's exposure to mass media has a significant effect on the use of health facilities for delivery. According to Kistiana, results from logistic regression analysis demonstrates that compared to those mothers who never or less than once in a month read, listened or watched any medium, mothers who were frequently exposed to any media had a 1.59 times greater chance of delivering their babies at health facilities (Kistiana, 2009). The same study in South India also explained that, mothers who exposed to mass media were more likely to receive assistance from health professionals than those who had less or no exposure, it varies in different states of the region (Navaneetham & Dharmalingam, 2000).

Religion: religion has influence the pregnant women in using services to treat and prevent maternal morbidity and mortality. A study from Bangladesh reveals that the use of skilled attendants at birth is higher among women from Hindu religion compared to women from Muslim religion (Anwar et al, 2008). On the other hand, studies in India claim that Muslim women are more likely to deliver with skilled assistance compared to women from Hindu religion (Thind et al, 2008, Navaneetham & Dharmalingam 2000).

Husband education: Findings from Indonesia explained the relation of husband education with delivery service utilization as, women with more educated husbands had higher odds ratio using skilled professional birth attendants than those whose husbands had less education. The probability of using skilled birth attendants increases by about 1.18 times for husbands with primary educational level than those women with no educated husbands' (Kistiana, 2009). Eyerusalem also revealed that women with partners who had a secondary or higher education had two times higher odds of delivering with professional assistance when compared to those with partners having no education (Eyerusalem, 2010).

2.3 OVERVIEW OF MODEL FAMILIES

Proper analysis of data is required in modeling the association between the response variable and the given set of covariates. Molenberghs & Verbeke broadly classified models in to two main model families (Molenberghs & Verbeke, 2005).

Marginal models: in which responses are modeled, marginalized overall other responses; the association structure is then typically captured using a set of association parameters, such as correlations, odds ratios, etc. Generalized estimating equation (GEE) and alternating logistic regression (ALR) are among marginal model family.

Cluster-specific models: the responses are assumed independent, given a collection of cluster-specific parameters. Generalized linear mixed model is one of subject specific family (Molenberghs & Verbeke, 2005). Based on the nature of sampling design and nature of data, some of the model families would be appropriate for this study is discussed as follow.

2.3.1 Generalized Estimating Equation (GEE)

According to Agresti, computationally simple alternative to maximum likelihood (ML) for clustered categorical data is a multivariate generalization of quasi likelihood. Rather than assuming a particular type of distribution for the response variable, this method only links each marginal mean to a linear predictor and provides a guess for the variance covariance structure of the response. The method uses the observed variability to help generate appropriate standard errors and called the GEE method because the estimates are solutions of generalized estimating equations. These equations are multivariate generalizations of the equations solved to find ML estimates for generalized linear models (Agresti, 2007).

Generalized estimating equations (GEE) models are a direct extension of basic quasi-likelihood theory from cross-sectional to repeated or otherwise correlated measurements. They estimate the parameters associated with the expected value of an individual's vector of binary responses and phrase the working assumptions about the association between pairs of outcomes in terms of marginal correlations (Molenberghs & Verbeke, 2005).

When we are mainly interested in first-order marginal mean parameters and pair wise interactions, a full likelihood procedure can be replaced by quasi-likelihood based methods (McCullagh and Nelder, 1989). In quasi-likelihood, the mean response is expressed as a parametric function of covariates, and the variance is assumed function of the mean up to possibly unknown scale parameters.

Wedderburn first noted that likelihood and quasi-likelihood theories coincide for exponential families and that the quasi-likelihood estimating equations provide consistent estimates of the regression parameters in any generalized linear model, even for choices of link and variance functions that do not correspond to exponential families (Wedderburn 1974). Consequently, Liang and Zeger proposed the method of generalized estimating equations (GEE) as an extension of GLM to accommodate correlated data using quasi-likelihood approach. Rather than assuming a particular distribution for the response, GEE method requires a correct specification of the mean as well as how the variance depends on the mean. One of the desirable properties of the GEE method is that it yields consistent and asymptotically normal solutions even with the misspecification of the covariance structure (Liang and Zeger, 1986).

In the methodology of generalized estimating equations, the user may impart a correlation structure that is often called a working correlation matrix. One often does not know what the true correlation is, hence, the term working correlation. Common correlation structures include; *Unspecified*: all correlations are to be independently estimated from the data, *Exchangeable*: all correlations within subjects are equal, *Independent*: all correlations are assumed to be zero (Myers *et al*, 2010). Because GEE does not have likelihood function, likelihood-ratio methods are not available for checking fit, comparing models, and conducting inference about parameters.

2.3.2 Alternating Logistic Regression (ALR)

Generalized estimating equation (GEE), allows estimation of first and second moment parameters in regression models for multivariate binary data. When association among the observation is importance and is measured using marginal odds ratios, the computations required will exclude the applications in studies with large clusters. An alternative approach that overcomes the computational limitations encountered in many problems is proposed what is called alternative logistic regression (Zeger *et*, 1993). As explained by Zeger *et al*, alternating logistic regression is reasonably efficient relative to GEE. In ALR, we estimate the association parameters by modeling the conditional distribution of one response given another.

Molenberghs & Verbeke also expressed ALR as extension of classical GEE, in the sense that precision estimates follow for both the parameters. However, unlike with GEE, no working assumptions about the third- and fourth-order odds ratios are required. The clever combination of a marginal and a conditional specification, addressing the third and fourth moments is avoided all together, which is strictly different from setting them equal to zero. This combination of marginal and conditional specification can be advantageous of ALR (Molenberghs & Verbeke, 2005).

2.3.3 Generalized Linear Mixed Model (GLMM)

Agresti explained that, generalized linear model (GLM) extend ordinary regression by allowing non-normal responses and a link function of the mean. The generalized linear mixed model is a further extension that permits random effects as well as fixed effects in the linear predictor (Agresti, 2007). Antonio & Beirlant defined GLMM as extend of GLM by allowing for random or cluster-specific effects in the linear predictor. These models are useful when the interest of the analyst lies in the individual response profiles rather than the marginal mean. The inclusion of random effects in the linear predictor reflects the idea that there is natural heterogeneity across subjects or clusters in some of their regression coefficients (Antonio & Beirlant, 2006).

According to McCulloch clarification, GLMM is very versatile in that they can handle non-normal data, nonlinear models, and a random effects covariance structure. This can be used to incorporate correlations in models, model the correlation structure, identify sensitive subjects and can be used to handle heterogeneous variances. The modeling process is relatively straightforward, requiring the following decisions: what is the distribution of the data, what is to be modeled, what are the factors, and are the factors fixed or random? This all makes GLMM attractive for use in modeling. Unfortunately, computing methods for much of the class of GLMM is an area of active research. No general-purpose software exists and, tests and confidence intervals are asymptotic and approximate (McCulloch, 1997).

Generalized the above explanation, GLMM is an extension to generalized linear model (GLM) that includes random effects in the linear predictor, giving an explicit probability model that explains the origin of the correlations. The resulting cluster-specific parameter estimates are suitable when the focus is on estimating the effect of changing one or more components of the predictor on a given individual.

The key problem in GLMM is maximization of the marginal likelihood, obtained by integrating out the random effects. In general, no analytic expressions are available for the integrals and numerical approximations are needed. There are large statistical literatures on various methods like approximation of the data, approximation of the Integral (Molenberghs & Verbeke, 2005).

To summarize, this brief literature review has shown the importance of a range of characteristics in determining maternal delivery service behavior. Some determinant covariates such as mother's education level, mother's age at birth, birth order, wealth index, religion, exposure to mass media and husbands education level are assessed, which are assumed to have positive or negative associations with the utilization of maternal delivery services. Some important model families like marginal models (GEE & ALR) and cluster specific model (GLMM) which are appropriate for analysis and the nature of the given data were assessed.

CHAPTER THREE

DATA AND METHODOLOGY

3.1 SOURCE OF DATA

The source of data for this study was the 2011 Ethiopia Demographic and Health Survey (EDHS), which is obtained from Central Statistical Agency (CSA). It is the third survey conducted in Ethiopia as part of the worldwide Demographic and Health Surveys project. The 2011 Ethiopian Demographic and Health Survey, was designed to provide estimates for the health and demographic variables of interest for the following domains. Ethiopia as a whole; urban and rural areas (each as a separate domain); and 11 geographic administrative regions (9 regions and 2 city administrations), namely: Tigray, Affar, Amhara, Oromiya, Somali, Benishangul-Gumuz, Southern Nations, Nationalities and Peoples (SNNP), Gambela and Harari regional states and two city administrations, that is, Addis Ababa and Dire Dawa.

The principal objective of the 2011 EDHS is to provide current and reliable data on fertility and family planning behavior, child mortality, adult and maternal mortality, children's nutritional status, use of maternal and child health services, knowledge of HIV/AIDS, and prevalence of HIV/AIDS and anemia.

3.1.1 Sampling Design of EDHS 2011

The 2007 Population and Housing Census, conducted by the CSA, provided the sampling frame from which the 2011 EDHS sample was drawn. Administratively, regions in Ethiopia are divided into zones, and zones, into administrative units called weredas. Each wereda is further subdivided into the lowest administrative unit, called kebele. During the 2007 Census, each kebele was subdivided into census enumeration areas (EAs) or clusters, which were convenient for the implementation of the census. The 2011 EDHS sample was selected using a stratified, two-stage cluster sampling design.

Clusters were the sampling units for the first stage. The sample included 624 clusters, 187 in urban areas and 437 in rural areas. Households comprised the second stage of sampling. In the second stage, a fixed number of 30 households were selected for each cluster. A complete listing of households was carried out in each of the selected clusters from September 2010 through January 2011.

The 2011 EDHS used three questionnaires: the Household Questionnaire, the Woman's Questionnaire, and the Man's Questionnaire. These questionnaires were adapted from model survey instruments developed for the measure DHS project to reflect the population and health issues relevant to Ethiopia. In addition to English, the questionnaires were translated into three major local languages-Amharigna, Oromiffa, and Tigrigna.

All women aged 15-49 and all men aged 15-59 were eligible for interview. In the interviewed households 17,385 eligible women were identified for individual interview; complete interviews were conducted for 16,515, yielding a response rate of 95%. In the survey, information on delivery care was collected from women who had at least one birth in the five years before the survey. The total numbers of women who had at least one birth in the last five years were 7758. There were 21cases in which information on the relevant variables was missing and these cases were excluded from the analysis. At the end, a total of 7737 cases from 594 clusters were included in the analysis.

3.1.2 Variables Considered in the Study

The response (dependant) and predictor variables that served for the estimation of parameters are defined as follows.

Response variable

In this study, the response variable is created from questions included in the maternal health component of the EDHS questionnaire. Although assistance during delivery is highly associated with place of delivery, it has been treated as a common variable because home deliveries can also be attended by health personnel. The primary outcome variable of this study was assistance during delivery, which defined as whether the mothers received assistance from a certified health professional (doctor, nurse or midwife).

This was coded as 1 if a mother received assistance from a health professional or trained health attendant and coded as 0 if otherwise.

Table 3.1 coding and explanation of response variable

Variable	Presentation of variable	Factor coding
Delivery service	Delivery	0= no any assistance 1= woman obtained services from any health professionals.

Predictor (explanatory) variables

The explanatory variables that would be included are explained as follow. The choice of these variables is guided by different literatures as the determinant factors of maternal delivery care service. These categories of the independent variables were coded starting from zero to make it appropriate for further analysis using different statistical models.

Table 3.1 Coding and description of explanatory variables

Variables	Explanation
Age of respondent (Age)	This variable refers to age of mother's at the time of the survey and has three categories ranging from 15-49 as 0 for 15-19, 1 for 20-39 & 2 for 40-49.
Educational status (Medul)	Educational status refers to the highest educational level mother's attained and it was categorized into three groups as no education, primary and secondary & higher and, categorized as 0 for No education, 1 for Primary & 2 for Secondary & above.
Mass Media (Massme)	Refers that the frequency of listening to radio, watching TV & reading Newspaper. Not at all means completely not exposed to any media, less than once a week means that expose to any media once in two, three or more weeks but not weekly and coded as 0 for Not atall,1 for less than once a week & 2 for at least once a week.
Educational level of Partner (Hedul)	Similar to educational level of the mothers and this was categorized into three groups as no education, primary and secondary plus higher education and the categories as 0 for No, education, 1 for Primary, 2 for Secondary & above

Work status (Workst)	In the survey, this was defined as if mother's has been working in any field other than household work in the seven days before the survey. This was classified as working or not working and categorized as 0 for Not working and 1 for Working
Place of residence (Residence)	This is a dichotomous variable (urban and rural) according to where the woman was living at the time of the survey as 0 for Rural and 1 for Urban
Religion (Religion)	Classification of this variable was developed according to previous literature by merging together the orthodox and catholic religion because mothers in the catholic group were few. The other categories are protestant, Muslim and other religion as 0 for Orthodox, 1 for Protestant, 2 for Muslim & 3 for others.
Wealth index (Wealth)	Measured by a composite score of several indicators of household possession. This was based on the questions about whether the household has items and facilities as piped water, toilet, type of floor used, electricity, radio, television and bicycle. Then according to the answer, each asset was given weight. Each household then was assigned a score according to each asset and the scores were summed for each household. It is coded as 0 for Poor,1 for Middle & 2 for Rich
Sex of household head(Sexhh)	This is classified as male or female. Based on the answer from the usual residents of the households on who the head of the household is and coded as 0 for Female and 1 for Male
Birth order (Brthord)	This refers to the rank of the child at birth. It has three categories starting from 1, 2-4, and 5+. For example, 1 refers to the first born child and coded as 0 for 1, 1 for 2-4 and 2 for 5+

3.2 METHOD OF DATA ANALYSIS

A range of techniques has been developed for analyzing data with categorical and clustered response variables. For this study, some extension of generalized linear models such as marginal models and cluster specific modes would be applied.

3.2.1 Generalized Linear Models (GLM)

Generalized linear models (GLMs) extend ordinary regression models to encompass non-normal response distributions and modeling functions of the mean (Agresti, 2002). Three components that specify a generalized linear model are *random component*, which identifies the response variable Y and its probability distribution; a *systematic component* specifies explanatory variables used in a linear predictor function; and a *link function* specifies the function of expected value of the response variable that the model equates to the systematic component. In general, GLM is a linear model for a transformed mean of a response variable that has distribution in the natural exponential family.

The Exponential Family

A random variable Y follows a distribution that belongs to the exponential family, if the density function is of the form

$$\mathrm{f}(y/\theta, \phi) = exp\{\phi^{-1}[y\theta - \psi(\theta)] + c(y, \phi)\} \tag{3.1}$$

, for a specific set of unknown parameters θ and ϕ, and for known functions $\psi(\cdot)$ and $c(\cdot,\cdot)$. The parameter θ is called the canonical parameter and represents the location while, ϕ is called the dispersion parameter and represents the scale parameter and for the Poisson and binomial distribution it is fixed to be one (Faraway, 2006). An important property of the GLM is the functional relation between mean and variance.

Generalized linear model assumes that the response variables are independent. In clustered data however, observations are usually taken from the same unit, and thus this information forms a cluster of correlated observations. For instance, in the EDHS the dependent variable (delivery care service) was measured once for each eligible mothers nested within clusters from each region.

3.2.2 Marginal Models

In clustered data, observations are usually taken from the same unit, and thus this information forms a cluster of correlated observations. Proper analysis of clustered data is required in modeling the association between the response variable and the given set of covariates. Marginal models are among the most statistical models widely used to model clustered or repeated data. The primary objective of marginal model is to analyze the population-averaged effects of the given factors in the study on the binary response variable of interest. This means that the covariates are directly related to the marginal expectations (Molenberghs & Verbeke, 2005). The marginal models fitted in this study would be included are Generalized Estimating Equations (GEE) and Alternating Logistic Regression (ALR).

3.2.2.1 Generalized Estimating Equations (GEE)

For binary data, a GEE approach is used to account for the correlation between responses of interest for subjects from the same cluster (Diggle et al., 1994). GEE is non-likelihood method that uses correlation to capture the association within clusters or subjects in terms of marginal correlations (Molenberghs & Verbeke, 2005). For clustered as well as repeated measured data, (Liang & Zeger, 1986) proposed GEE which require only the correct specification of the univariate marginal distributions provided one is willing to adopt "working" assumptions about the correlation structure. The "working" assumptions as proposed by Liang and Zeger, included independence, unstructured, exchangeable and auto-regressive AR (1). Independence and exchangeable working assumptions can be used in virtually all applications, whether longitudinal, clustered, multivariate, or otherwise correlated. Auto regressive AR (1) and unstructured correlation structures are less relevant for clustered data, studies with unequally spaced measurements or sequences with differing lengths (Molenberghs and Verbeke, 2005).

Let $Y_j = (y_{j1}, \dots\dots y_{jn_j})'$ be the response values of observations from $\boldsymbol{j^{th}}$ cluster**,** for $j = 1,2, \dots, m$ follows a binomial distribution i.e $\boldsymbol{Y_j \sim Bino(n_j, \pi_j)}$ that belongs to the exponential family with the density function of the form (3.1). Then, to model the relation between the response and covariates, one can use a regression model similar to the generalized linear models given by:

$$\boldsymbol{g(\pi_j) = logit(\pi_j) = X'_j\beta} \qquad (\mathbf{3.2})$$

Where, $g(\pi_j)$= logit link function, $\boldsymbol{X_j}$= (n_j x p) dimensional vector of known covariates, $\boldsymbol{\beta}$= (1 x p) dimensional vector of unknown fixed regression parameter to be estimated and $\boldsymbol{E(Y_j) = \pi_j}$ *is* expected values of the response variable from j^{th} cluster.

3.2.2.1.1 Parameter Estimation for GEE

As previously expressed GEE is not likelihood approach, rather it is quasi-likelihood based and estimates $\hat{\beta}$ by solving estimating equations which consist of the working covariance matrix V_j. The score equation used to estimate the marginal regression parameters while accounting for the correlation structure is given by:

$$\boldsymbol{S(\beta) = \sum_{j=1}^{m} \frac{\partial \pi_j}{\partial \beta'} \left[A_j^{1/2} R_j A_j^{1/2}\right]^{-1} (Y_j - \pi_j) = 0} \qquad (\mathbf{3.3})$$

Where $\boldsymbol{R_j}$ is working correlation matrix, and the covariance matrix of $\boldsymbol{Y_j}$ is decomposed in to $\boldsymbol{A_j^{1/2} R_j A_j^{1/2}}$ with $\boldsymbol{A_j}$ the matrix with the marginal variances on the main diagonal and zeros elsewhere and $\boldsymbol{Y_j}$ is multivariate vector of asymptotically normal response variables with mean vector $\boldsymbol{\pi_j}$ i.e $\boldsymbol{Y_j \cong N(X_j\beta, V_j)}$. An advantage of the GEE approach is that it yields a consistent estimator of $\hat{\beta}$, even when the working correlation matrix $\boldsymbol{R_j}$ is misspecified. However, severe misspecification of working correlation may seriously affect the efficiency of the GEE estimators (Molenberghs & Verbeke, 2005).

3.2.2.2 Alternating Logistic Regression (ALR) Model

This method is very similar to that of GEE, in that they are both quasi-likelihood based and they account for dependency in the data. However, unlike GEE which measures the association among the observed data through the correlation structure; Alternating logistic regression (ALR) measures this association using the odds ratio, which is interpretable and more applicable for binary data. ALR extends beyond classical GEE in the sense that precision estimates follow for both the regression parameters $\boldsymbol{\beta}$ and the association parameters $\boldsymbol{\alpha}$**.** Moreover, with ALR inferences can be made, not only about marginal parameters but also about pair wise associations between subjects as well (Molenberghs & Verbeke, 2005).

For cluster $j = 1, 2, \ldots\ldots, m$, let $Y_j = \left(Y_{j1}, \ldots . Y_{jn_j}\right)'$ be a $n_j \times 1$ response vector with mean $E\left(Y_j\right) = \pi_j$ and let ψ_{jkl} be the odds ratio between responses Y_{jk} and Y_{jl} $(1 \leq k \leq l \leq n_j)$ defined by

$$\boldsymbol{\psi_{jkl}} = \frac{\boldsymbol{P\left(Y_{jk} = 1, Y_{jl} = 1\right)P\left(Y_{jk} = 0, Y_{jl} = 0\right)}}{\boldsymbol{P\left(Y_{jk} = 1, Y_{jl} = 0\right)P\left(Y_{jk} = 0, Y_{jl} = 1\right)}} \qquad (3.3)$$

$j = 1,2, \ldots \ldots m$, $\quad k, l = 1,2, \ldots \ldots \ldots n_j, where$, Y_{jk} and Y_{jl} represents the response values for mothers k and l respectively from the same cluster. Let γ_{jkl} be the log odds ratio between outcomes Y_{jk} and Y_{jl}, let $\pi_{jk} = P\left(Y_{jk} = 1\right)$ and $v_{jkl} = P\left(Y_{jk} = 1, Y_{jl} = 1\right)$,then the association of the two responses (Zeger *et al*, 1993) is defined as:

$$\boldsymbol{logitP\left(Y_{jk} = 1/Y_{jl} = y_{jl}\right) = \gamma_{jkl}y_{jl} + log\left(\frac{\pi_{jk} - v_{jkl}}{1 - \pi_{jk} - \pi_{jl} + v_{jkl}}\right)} \qquad (3.4)$$

Assume $\boldsymbol{\gamma_{jkl} = \alpha}$. Then the pairwise log odds ratio α is the regression coefficient in logistic regression of Y_{jk} on Y_{jl} as long as the second term on the right-hand side in (3.4) is used as an offset. Generally $\boldsymbol{log(\psi_{jkl}) = \gamma_{jkl} = z'_{jkl}\alpha, where\ z_{jkl}}$ is a q×1 vector of covariates which specifies the form of the association between Y_{jk} and Y_{jl}.

3.2.2.2.1 Parameter Estimation of ALR

Since ALR also not maximum likelihood approach like GEE, parameter estimation is based on the score equation of the approximate likelihood that is based on quasi likelihood approximation. Let $\boldsymbol{\mu_j}$ be a vector with elements $\mu_{jkl} = E\left(Y_{jk}/Y_{jl} = y_{jl}\right)$ and let $\boldsymbol{R_j}$ be the vector of residual with elements $R_{jkl} = Y_{jk} - E\left(Y_{jk}/Y_{jl} = y_{jl}\right) = Y_{jk} - \mu_{jkl}$.

Let $\boldsymbol{S_j}$ a vector of diagonal matrix with diagonal element $\boldsymbol{\mu_{jkl}\left(1 - \mu_{jkl}\right)}$ and let $\boldsymbol{W_j}$ denote matrix $\boldsymbol{\frac{\partial \mu_j}{\partial \alpha}}$. Finally, let $\boldsymbol{A_j = Y_j - \pi_j}$, $\boldsymbol{B_j = cov\left(Y_j\right)}$, $\quad \boldsymbol{C_j = \frac{\partial \pi_j}{\partial \beta}}$.

Then the alternating logistic regression parameter $\boldsymbol{\delta}$=$(\boldsymbol{\beta}, \boldsymbol{\alpha})$ is the simultaneous solution of the following unbiased estimating equations (Zeger *et*, 1993).

$$\boldsymbol{U_\beta = \sum_{j=1}^{m} C'_j B_j^{-1} A_j = 0} \qquad (3.5)$$

$$\boldsymbol{U_\alpha = \sum_{j=1}^{m} W'_j S_j^{-1} R_j = 0} \qquad (3.6)$$

Estimating equation 3.5 and 3.6 are solving for β and α by using Gauss-Seidel procedure algorithm. ALR is computationally feasible for very large cluster.

3.2.2.3 Model Building for Marginal Models

Model selection is an important issue in almost any practical data analysis. A common problem is variable selection in regression given a large group of covariates (including some higher order terms) one needs to select a subset to be included in the regression model. Model selection is data analysis strategy, which leads to a search of best model. With this, we mean selecting the best subset of the covariates from the available covariates in the data.

3.2.2.3.1 Variable Selection Technique

To select significant variables, firstly under the GEE, model building strategy started by fitting a model containing all possible covariates in the data. This was done by considering two working correlation assumptions (exchangeable and independence). In order to select the important factors related to the response variable, the backward selection procedure was used. The strategy is called backward because we are working backward from our largest starting model to a smaller final model. In this case, the procedure is used to remove covariates with non-significant p-values. This means that variables that did not contribute to the model based on the highest p-value was eliminated sequentially and each time a new model with the remaining covariates was refitted, until we remained with covariates necessary for answering our research question. Finally, the two models were compared using model comparison techniques. Additionally, using the same procedures, an ALR model, which provides information about pair wise association of observations between two different individuals within the same cluster, was fitted. It turned out that the model with selected covariates is found to be the most parsimonious model.

3.2.3.3.2 Model Comparison Technique

Quasi-Information Criterion (QIC)

In a condition, when the likelihood function cannot be fully specified, such as in the GEE case, the Akaike's Information Criterion (AIC) cannot be directly applied to select either the optimal set of explanatory variables or correlation matrix. As an alternative, one can use the modified Akaike's Information Criterion called Quasi Information Criteria (QIC), which is based on the quasi-likelihood function (Pan, 2001). QIC is derived from the AIC and conceptually similar. The quasi-likelihood function takes the following form (McCullagh & Nelder, 1989)

$$\mathbf{Q}(\boldsymbol{\pi}) = \int_{y}^{\pi} \frac{\boldsymbol{y-t}}{\boldsymbol{\Phi}\mathbf{v}(\mathbf{t})} \boldsymbol{dt}$$

Where $\pi = E(y), v(y) = \Phi v(\pi)$, and Φ being the dispersion parameter.

An equation for the QIC is

$\mathbf{QIC} = -\mathbf{2Q}\,(\hat{\boldsymbol{\pi}}, \mathbf{I}) + \mathbf{2trace}[(\Omega_{\boldsymbol{I}}{}^{-1}\hat{\boldsymbol{V}}_{\boldsymbol{R}}]$ Where **I** represent the independent correlation structure (diagonal matrix) and ***R*** is the specified working correlation structure. The *p*-dimensional matrices $\boldsymbol{\Omega_I}^{-\mathbf{1}}$**and** $\hat{\boldsymbol{V}}_{\boldsymbol{R}}$ are variance estimators of the regression coefficients under the correlation structure **I** and **R** respectively. The QIC value is computed based on the quasi-likelihood estimate $\hat{\boldsymbol{\pi}}$ and is used to select the candidate explanatory variables. The model with the smallest QIC value for all correlation structures is considered as the best candidate model.

The generalized Wald test: is used to compare models with different subsets of the regression parameters, i.e to select the candidate covariates. That is, one can use the generalized Wald tests to test the joint null hypothesis that a set of regression parameters $\boldsymbol{\beta}$s are equal to zero. In general, for any matrix **L** a test for hypothesis can be written as follows

$H_0\!: L\boldsymbol{\beta} = \mathbf{0}$ ***versus*** $H_1 : L\boldsymbol{\beta} \neq \mathbf{0}$**,** Where **L** is a **p x q** indicator matrix of ones and zeros. Here, ***p*** is equal to the number of parameters in the full model (including the intercept) and ***q*** equals the number of parameters in the generalized Wald test (that is, the difference in parameters between the full and reduced model). The Wald statistic is a quadratic form defined as follows

$\boldsymbol{W^2}_{\boldsymbol{stat}} = \hat{\boldsymbol{\beta}}^{\boldsymbol{t}}\ \mathbf{L}^{\boldsymbol{t}}(\mathbf{L}\,\mathbf{Var}(\boldsymbol{\beta})\,\mathbf{L}^{\boldsymbol{t}})^{-\mathbf{1}}\mathbf{L}\hat{\boldsymbol{\beta}}$. It is distributed as $x^2$with *q* degrees of freedom under the null hypothesis.

In addition to select the appropriate working correlation structure, the two models with exchangeable and independence working correlation were compared via their naïve (model based) and robust (empirical) standard error estimates and the one with the closest empirical and model based standard error estimates was preferred (Molenberghs & Verbeke, 2005). Moreover, unless one expects dramatic differences among the correlations, using the exchangeable working correlation structure is recommended (Agresti, 2007).

3.2.3 Cluster Specific (Subject Specific) Models

When interest is in the marginal or population-averaged models to analysis the relationships of the covariates to the dependent variable for an entire population, marginal models as discussed in previous section are preferred. However, in most biomedical and biological data problems, interest often lies in understanding the response of individual patient characteristics and how this response is influenced by a given set of possible covariates (Myers et al.,2010). This proves even to be essential when individual interventions may be necessary. Cluster specific models are useful in such cases. Cluster specific models differ from the marginal models by inclusion of parameters that are specific to clusters or subjects within a population. Consequently, random effects are directly used in modeling the random variation in the dependent variable at different levels of the data.

3.2.3.1 Generalized Linear Mixed Model (GLMM)

Generalized linear models (GLM) is one parts of subject specific models which extends ordinary regression by allowing non-normal responses and a link function of the mean. The generalized linear mixed model is a further extension that permits random effects as well as fixed effects in the linear predictor (Agresti, 2002).

Let y_{ij} denote the response of i^{th}individual mother from j^{th}cluster where $i = 1,2 \ldots \ldots \ldots \ldots n_j$ and Y_j the n_j dimensional vector of all measurements available for cluster j. Let $\boldsymbol{f(b_j/D)}$ be the density of the $\boldsymbol{N(0, D)}$ distribution for the random effects $\boldsymbol{b_j}$. Assumed conditionally on q-dimensional random effects $\boldsymbol{b_j}$ to be drawn independently from $\boldsymbol{N(0, D)}$, the outcomes $\boldsymbol{y_{ij}}$ of $\boldsymbol{Y_j}$ are independent with the density of the form

$$\boldsymbol{f_j(y_{ij}/b_{j,}\beta, \phi) = exp\{\phi^{-1}[y_{ij}\theta_{ij} - \psi(\theta_{ij})] + c(y_{ij}, \phi\}}, \qquad (\mathbf{3.7})$$

Then the generalized linear mixed model (Molenberghs and Verbeke, 2005); with logit link is defined as

$$\boldsymbol{logit(\pi_{ij}) = X'_{ij}\beta + Z'_{ij}b_j, \quad j = 1, 2, \ldots, m} \qquad (\mathbf{3.8})$$

Where, $\boldsymbol{E(Y_{ij}/b_j) = \pi_{ij}}$, is the mean response vector conditional on the random effects $\boldsymbol{b_j}$, for mothers in cluster j and, X_{ij} and $\boldsymbol{Z_{ij}}$ are *p*-dimensional and *q*-dimensional vectors of known covariate values. The random effects $\boldsymbol{b_j}$ are assumed to follow a multivariate normal distribution with mean **0** and covariance matrix ***D.***

3.2.3.1.1 Parameter Estimation for GLMM

Random-effects models can be fitted by maximization of the marginal likelihood, obtained by integrating out the random effects. Such likelihood may involve high-dimensional integrals that cannot be evaluated analytically. The likelihood of the data expressed as a function of unknown parameters is

$$\mathbf{L}(\boldsymbol{\beta}, \boldsymbol{D}, \phi) = \prod_{j=1}^{m} \boldsymbol{f_j}(\boldsymbol{Y_j}/\boldsymbol{\beta}, \boldsymbol{D}, \phi) = \prod_{j=1}^{m} \int \prod_{j=1}^{n_j} \boldsymbol{f_{ij}}(\boldsymbol{Y_{ij}}/\boldsymbol{b_j}, \boldsymbol{D}, \phi) \boldsymbol{f}(\boldsymbol{b_j}/\boldsymbol{D}) \boldsymbol{d}\boldsymbol{b_j} \qquad (\mathbf{3.9})$$

It is the integral over the unobserved random effects of the joint distribution of the data and random effects. The problem in maximizing (3.9) is the presence of m integrals over the q-dimensional random effects b_j. With Gaussian data, the integral has a closed form solution and relatively simple methods exist for maximizing the likelihood or restricted likelihood. With non-linear models, numerical techniques are needed. The Laplace method (Molenberghs & Verbeke, 2005) has been designed to approximate integrals of the form:

$$\boldsymbol{I} = \int \boldsymbol{e}^{\boldsymbol{Q}(\boldsymbol{b})} \boldsymbol{db} \qquad (\mathbf{3.10})$$

Where $Q(b)$ is a known, unimodal, and bounded function of a q-dimensional variable ***b*.** Let $\widehat{\boldsymbol{b}}$ be the value of ***b*** for which Q is maximized. Then the second order Taylor expansion of $Q(b)$ is the form

$$\boldsymbol{Q}(\boldsymbol{b}) \approx \boldsymbol{Q}(\widehat{\boldsymbol{b}}) + \frac{1}{2}(\boldsymbol{b} - \widehat{\boldsymbol{b}})' \boldsymbol{Q}''(\widehat{\boldsymbol{b}})(\boldsymbol{b} - \widehat{\boldsymbol{b}}) \qquad (\mathbf{3.11})$$

where, $\boldsymbol{Q}''(\widehat{\boldsymbol{b}})$ is the matrix of second-order derivative of Q, evaluated at $\widehat{\boldsymbol{b}}$**.** Replacing$\boldsymbol{Q}(\boldsymbol{b})$ in (3.10) by its approximation in (3.11), we obtain

$$\boldsymbol{I} \approx (\mathbf{2}\boldsymbol{\pi})^{\boldsymbol{q}/\mathbf{2}} \left|-\boldsymbol{Q}\left(\widehat{b}\right)\right|^{-\mathbf{1}/\mathbf{2}} \boldsymbol{e}^{\boldsymbol{Q}\left(\widehat{b}\right)} \qquad (\mathbf{3.12})$$

Clearly, each integral in (3.9) is proportional to an integral of the form (3.10) for functions $Q(b)$ given by

$$\boldsymbol{Q}(\boldsymbol{b}) = \phi^{-\mathbf{1}} \sum_{i=1}^{n_j} [\boldsymbol{y_{ij}}(\boldsymbol{x'_{ij}}\boldsymbol{\beta} + \boldsymbol{z'_{ij}}\boldsymbol{b}) - \boldsymbol{\psi}(\boldsymbol{x'_{ij}}\boldsymbol{\beta} + \boldsymbol{z'_{ij}}\boldsymbol{b})] - \frac{\mathbf{1}}{\mathbf{2}} \boldsymbol{b'}\boldsymbol{D}^{-\mathbf{1}}$$

This is called the Laplace's method or approximation of integrands. Note that the mode $\widehat{\boldsymbol{b}}$ of Q depends on the unknown parameters β, ϕ, and **D**, such that in each iteration of the numerical maximization of the likelihood, $\hat{\boldsymbol{b}}$ will be recalculated conditionally on the current values for the estimates for these parameter.

3.2.3.1.2 Model Building for GLMM

A different approach to account for clustering is by using random components such as random intercepts. Under the GLMM, model building was begun by adoption of the marginal model covariates. Additionally, the model also included the random effects ,in this case, random intercepts to address the between and within-regional heterogeneity. These were introduced in the generalized linear mixed model due to the fact that, the probability of using delivery service possibly varies for individuals within the same regions as well as individuals in different regions. Variable selection procedure for GLMM is similar with marginal model previously explained.

3.2.3.1.3 Model Comparison in GLMM

This study will be used Likelihood ratio test and Information criteria to select the best model based on the values of asymptotic estimations.

Likelihood Ratio Test

In order to decide on the best of the two random effects models, two models will be fitted, one with the two random intercepts (between and within regional variations) and another with one random intercept (within regional variation). One can use the approximate restricted maximum likelihood ratio test (LRT) to compare these two models (Myers et al., 2010). Let $\boldsymbol{LR_{full} = -2\log likelihood}$ value for the full model and $\boldsymbol{LR_{redu} = -2loglikelihood}$ value for reduced model. Then, the likelihood ratio test statistic, is given by

$$\boldsymbol{\lambda = LR_{full} - LR_{redu}}$$

The asymptotic null distribution of the likelihood ratio test statistic $\boldsymbol{\lambda}$, is a chi-square distribution with degrees of freedom equal to the difference between the numbers of parameters in the two models.

Akaike's information criterion (AIC)

AIC is a measure of goodness of fit of an estimated statistical model. It is not a test on the model in the sense of hypothesis testing; rather it is a tool for model selection. The AIC penalizes the likelihood by the number of covariance parameters in the model, therefore

$$\boldsymbol{AIC = -2\log(L) + 2p}$$

Where, L is the maximized value likelihood function for the estimated model and p is the number of parameters in the model. The model with the lowest AIC value is preferable.

3.2.3.1.4 Model Checking Technique

In GLMM, it is assumed that the random effects are normally distributed and uncorrelated with the error term. Normality of the random effects is assessed using normal plot of each random effect. Normal Q-Q plot of estimated random effects is an important method for checking the normality (Myers et al., 2010).

CHAPTER FOUR
ANALYSIS AND DISCUSSION

4.1 SUMMARY OF DESCRIPTIVE STATISTICS

Table 4.1 presents basic descriptive information that summarizes the associations between the determinant factors and delivery service utilization of mothers. The total of 7737 mothers from nine regional states and two city administrations in Ethiopia were eligible for this study. Among these eligible mothers, 1328 (17.2%) mothers were getting delivery service from health professionals where as 6409 (82.8%) are not get access of delivery service.

The proportion of delivery service is slightly larger (18.4%) for adolescent age groups than the youth mothers (14.7%). The delivery service is less (7.4%) for older age compared to middle age and youth mothers. There is also a remarkable variation of delivery service utilization due to place of residence of mothers. The proportion of getting delivery assistance for urban mothers is (63.4%) and who living in rural area is (6.6%) only.

Educational level of mothers has increasing proportion to delivery service utilization. The proportion of delivery service is (7.4%) for non educated mothers, (26.6%) for primary educated mother and (81%) for mothers whose education level is secondary and above. As compared to followers of the Coptic Orthodox religion who use the delivery services (25.7%), Protestant mothers are less to use delivery care service (14.6%). On the other hand, Muslim mothers tend to occupy an intermediate position in using delivery care services (11.9%) and other religion follower mother take the least position in utilization of delivery care services (7.9%).

Delivery care services coverage is also associated with sex of household head and wealth quintile. Mother's with male household head uses delivery care services less percentage (15.3%) than mothers with female household head (24.5%). Mother's with the lowest wealth index have less proportion to use health personal assistance during their most recent birth than mothers in higher wealth index. Mothers whose household wealth index is poor use delivery care services (3.4%). Middle wealth index mothers use delivery care services of (18.8%) and (37.4%), for mothers with rich wealth index. The proportion of mothers using delivery assistance service for the first birth is relatively higher (36.2%) than subsequent birth orders (17.9%) for birth orders two to four and (6.5%) for five and higher birth order.

Table 4.1 Summary of descriptive statistics for delivery assistance service use

Variables	**Levels**	**Assistance during delivery (%)**		**Total**
		No	Yes	
Age	15-19	354(85.3)	61(14.7)	415
	20-39	5365(81.6)	1212(18.4)	6577
	40-49	690(92.6)	55(7.4)	745
Residence	Urban	524(36.6)	909(63.4)	1433
	Rural	5885(93.4)	419(6.6)	6304
Medul	No education	4782(92.6)	380(7.4)	5162
	Primary	1535(73.4)	555(26.6)	2090
	Secondary & above	92(19.0)	393(81.0)	485
Religion	Orthodox	2001(74.3)	693(25.7)	2964
	Protestant	1380(85.4)	236(14.6)	1616
	Muslim	2794(88.1)	379(11.9)	3173
	Others	234(92.1)	20(7.9)	254
Sexhh	Female	1180(75.5)	382(24.5)	1562
	Male	5229(84.7)	946(15.3)	6175
Wealth	Poor	3023(96.6)	106(3.4)	3129
	Middle	2186(82.5)	464(18.8)	2650
	Rich	1200(61.3)	758(37.4)	1958
Brthord	1	940(63.8)	534(36.2)	1474
	2-4	2774(82.1)	606(17.9)	3380
	≥5	2695(93.5)	188(6.5)	2883
Hedul	No education	3655(94.1)	231(5.9)	3886
	Primary	2296(81.0)	538(19.0)	2834
	Secondary & above	458(45.0)	559(55.0)	1017
Workst	Non Employment	4546(85.8)	753(14.2)	5299
	Employment	1863(76.4)	575(23.6)	2438
Massme	Not at all	3224(94.7)	185(5.3)	3509
	Less than a week	1398(89.9)	157(10.1)	1555
	At least once a week	1687(63.1)	986(36.9)	2673
	Total	6409(82.8)	1328(17.2)	7737

Mothers whose husband's had secondary and higher educational level uses delivery care service highly (55.0%) than the primary educational level uses (19.0%) delivery care service. Mother's whose husband had no education uses delivery care service only (5.9%).

With regard to work status, the percentage of employed mothers were higher to use delivery care services (23.6%) than mothers who are not employed use delivery care services (14.2%). Delivery care service also has an increasing proportion to mass media. Mothers follow a mass media at least once in a week has higher proportion (36.9%) to use delivery service than that of follow less than a week that use the delivery service (10.1%). Mothers not follow mass media at all are uses delivery service only (5.3%).

4.2 STATISTICAL ANALYSIS OF MARGINAL MODELS

In this section, the delivery care service data was analyzed by using marginal models including generalized estimating equation and alternating logistic regression models.

4.2.1 Analysis of Generalized Estimating Equations (GEE)

In the methodology that is termed generalized estimating equations, the user may impart a correlation structure that is often called a working correlation matrix. Before selecting the correct correlation structure, consider the model building strategy (variable selection).

Under the GEE, model building strategy is started by fitting a model containing all possible covariates in the data. This was done by considering two different working correlation assumptions (exchangeable and independence). In order to select the important factors related to delivery service, the backward selection procedure was used. The full logit model for the probability of getting delivery service of i^{th} mother from j^{th} cluster (π_{ij}) was fitted as:

$$\begin{aligned} logit(\pi_{ij}) = \beta_0 &+ \beta_1 Age_1 + \beta_2 Age_2 + \beta_3 Residence_U + \beta_4 Medul_P + \beta_5 Medul_{S+} \\ &+ \beta_6 Religion_{Pr} + \beta_7 Religion_{Mu} + \beta_8 Religion_O + \beta_9 Sexhh_M + \beta_{10} Wealth_M \\ &+ \beta_{11} Wealth_R + \beta_{12} Brtord_{2+} + \beta_{13} Brtord_{5+} + \beta_{14} Hedul_P + \beta_{15} Hedul_{S+} \\ &+ \beta_{16} Workst_E + \beta_{17} Massme_L + \beta_{18} Massme_{At} \end{aligned}$$

The subscripts in each covariates are defined as, *1= 20-39, 2 = 40-49, U=Urban, P = Primary, S+ = Secondary and above, Pr = Protestant, Mu = Muslim, O = Other, M = male, Md = Middle, Ri= Rich, 2+ = 2-4, 5+ = ≥ 5, L = Less than once a week, At = At least once a week*.

After fitting the model, covariates with the largest p-value of Wald test is removed and refitted the model with the rest of the covariates sequentially. Then, work status of mothers and sex of household head are the covariates excluded from the model; with Wald test p-value for the given covariates are large (*P-value > 0.05*) which is found in the appendix. The QIC values of full model and reduced models are 4100.0609 (which is found in appendix) and 4098.4137 respectively. Then it turned out that the model with age, type of place of residence, mothers' education level, religion, wealth index, birth order, husband education level and exposure to mass media as covariates was the most parsimonious model.

Table 4.2 Empirical and model based standard errors for two proposed working correlations

	Exchangeable			**Independent**		
Coeff.	Estimates	***Model based S.E***	***Empirical S.E***	Estimates	***Model based S.E***	***Empirical S.E***
$\boldsymbol{\beta_0}$	-3.2994	***0.2265***	***0.2402***	-3.5073	*0.2177*	*0.2528*
$\boldsymbol{\beta_1}$	0.5085	***0.1835***	***0.2029***	0.5489	*0.1865*	*0.2064*
$\boldsymbol{\beta_2}$	0.5516	***0.2542***	***0.2556***	0.5772	*0.2642*	*0.2785*
$\boldsymbol{\beta_3}$	2.3172	***0.1133***	***0.1364***	2.1519	*0.0920*	*0.1406*
$\boldsymbol{\beta_4}$	0.4932	***0.0961***	***0.0987***	0.5097	*0.0968*	*0.1049*
$\boldsymbol{\beta_5}$	1.3755	***0.1673***	***0.1665***	1.5100	*0.1701*	*0.1794*
$\boldsymbol{\beta_6}$	-0.1981	***0.1215***	***0.1380***	-0.2924	*0.1085*	*0.1516*
$\boldsymbol{\beta_7}$	-0.6483	***0.1148***	***0.1260***	-0.6454	*0.0999*	*0.1213*
$\boldsymbol{\beta_8}$	-0.3940	*0.2886*	*0.2654*	-0.4465	***0.2814***	***0.2758***
$\boldsymbol{\beta_9}$	0.3194	***0.1301***	***0.1315***	0.4538	*0.1290*	*0.1566*
$\boldsymbol{\beta_{10}}$	0.8358	***0.1347***	***0.1411***	1.1247	*0.1297*	*0.1560*
$\boldsymbol{\beta_{11}}$	-0.7237	***0.1024***	***0.1083***	-0.7597	*0.1039*	*0.1122*
$\boldsymbol{\beta_{12}}$	-0.9846	***0.1275***	***0.1278***	-1.1242	*0.1307*	*0.1394*
$\boldsymbol{\beta_{13}}$	0.3190	***0.1008***	***0.1015***	0.3619	*0.1030*	*0.1129*
$\boldsymbol{\beta_{14}}$	0.6790	***0.1311***	***0.1228***	0.7796	*0.1321*	*0.1374*
$\boldsymbol{\beta_{15}}$	0.4241	***0.1189***	***0.1251***	0.4932	*0.1213*	*0.1416*
$\boldsymbol{\beta_{16}}$	0.7038	***0.1085***	***0.1082***	0.7591	*0.1084*	*0.1216*

Finally, as a customary, comparison of empirical and model based standard errors for the parameter estimates obtained based on the given working correlation assumptions (in this study exchangeable and independence) was performed using selected covariates. The correlation structure with the model based and empirical standard errors are closest to each other, is referred to be the best assumption correlation structure.

Moreover, since no dramatic differences among the correlations, using the exchangeable working correlation structure is recommended. In addition, the empirically corrected standard errors for exchangeable correlation structure are somewhat smaller than their counterpart under the independence assumptions.

Then, from table 4.2, exchangeable working correlation assumption was found to be plausible since the two standard errors were closer each other with ($\boldsymbol{\alpha}$ = 0.068).

Therefore, the final proposed generalized estimating equation model for delivery care service utilization of mothers was given as:

$$logit(\pi_{ij}) = \beta_0 + \beta_1 Age_1 + \beta_2 Age_2 + \beta_3 Residence_U + \beta_4 Medul_P + \beta_5 Medul_{S+} + \beta_6 Religion_{Pr} + \beta_7 Religion_{Mu} + \beta_8 Religion_O + \beta_9 Wealth_M + \beta_{10} Wealth_R + \beta_{11} Brtord_{2+} + \beta_{12} Brtord_{5+} + \beta_{13} Hedul_P + \beta_{14} Hedul_{S+} + \beta_{15} Massme_L + \beta_{16} Massme_{At}$$

Parameter estimates and their corresponding empirically corrected standard errors alongside the p-values from the final GEE model (model 4.2) are presented at table 4.3

Table 4.3 Parameter estimates (empirically corrected standard errors) for GEE

Effects	**Level**	***Parameter***	**Estimates (s.e)**	**Log 95% conf.int**	***P-value***
Intercept		β_0	*-3.2994(0.2402)*	*(-3.7702,-2.8285)*	*<.0001*
Age	15-19 (Ref)		.	.	.
	20-39	β_1	*0.5085(0.2029)*	*(0.1109,0.9061)*	*0.0122*
	40-49	β_2	*0.5516(0.2556)*	*(0.0505,1.0526)*	*0.0310*
Residence	Rural (Ref.)		.	.	.
	Urban	β_3	*2.3172(0.1364)*	*(2.0499,2.5845)*	*<.0001*
Medul	No education (Ref)		.	.	.
	Primary	β_4	*0.4932(0.0987)*	*(0.2996,0.6867)*	*<.0001*
	Secondary & above	β_5	*1.3755(0.1665)*	*(1.0493, 1.7018)*	*<.0001*
Religion	Orthodox (Ref)		.	.	.
	Protestant	β_6	*-0.1981(0.1380)*	*(-0.4686, 0.0724)*	*0.1512*
	Muslim	β_7	*-0.6483(0.1260)*	*(-0.8952, -0.4013)*	*<.0001*
	Other	β_8	*-0.3940(0.2654)*	*(-0.9142, 0.1261)*	*0.1376*
Wealth	Poor (Ref)		.	.	.
	Middle	β_9	*0.3194(0.1315)*	*(0.0616, 0.5771)*	*0.0152*
	Rich	β_{10}	*0.8358(0.1411)*	*(0.5592, 1.1123)*	*<.0001*
Brthord	1 (Ref)		.	.	.
	2-4	β_{11}	*-0.7237(0.1083)*	*(-0.9359, -0.5114)*	*<.0001*
	≥5	β_{12}	*-0.9846(0.1278)*	*(-1.2352, -0.7341)*	*<.0001*
Hedul	No education (Ref)		.	.	.
	Primary	β_{13}	*0.3190(0.1015)*	*(0.1200, 0.5179)*	*0.0017*
	Secondary & above	β_{14}	*0.6790(0.1228)*	*(0.4483, 0.9197)*	*<.0001*
Massme	Not at all (Ref)		.	.	.
	Less than once week	β_{15}	*0.4241(0.1251)*	*(0.1790,0.6692)*	*0.0007*
	At least once a week	β_{16}	*0.7038(0.1082)*	*(0.4917,0.9159)*	*<.0001*
	QIC		*4098.4137*		

4.2.2 Analysis of Alternating Logistic Regression Model (ALR)

Model building for ALR is follows the same procedure in GEE model building strategy. First ALR model was fitted using all proposed covariates. Then the covariate with the large p-value is removed. Work status of mother and sex of household head are removed covariates with Wald test (p-value > 0.05). The QIC values of both saturated and reduced models are given by 4125.6687 (found in appendix) and 2123.1463 respectively. Therefore, the reduced model with the rest of eight covariates was considered as the best candidate model. Using the selected covariates and the association parameter **α**, alternating logistic regression (ALR) model that provides information about pair wise association of observations between two different individuals within the same cluster was fitted. Therefore, the final proposed ALR model included the association parameter for delivery care service utilization of mothers is given at the next page as:

$$logit(\pi_{ij}) = \beta_0 + \alpha + \beta_1 Age_1 + \beta_2 Age_2 + \beta_3 Residence_U + \beta_4 Medul_P + \beta_5 Medul_{S+} + \beta_6 Religion_{Pr} + \beta_7 Religion_{Mu} + \beta_8 Religion_O + \beta_9 Wealth_M + \beta_{10} Wealth_R + \beta_{11} Brtord_{2+} + \beta_{12} Brtord_{5+} + \beta_{13} Hedul_P + \beta_{14} Hedul_{S+} + \beta_{15} Massme_L + \beta_{16} Massme_{At}$$

Parameter estimates and their corresponding empirically corrected standard errors alongside the p-values from the final ALR model are presented in table 4.4.

4.2.3 Comparison of GEE and ALR Models

Since the likelihood function does not fully specified in marginal models, model comparison is based on quasi likelihood criteria (QIC) which is the modified AIC criteria.

From table 4.3 and table 4.4, we found that the QIC values are *4098.4137* and *4123.1463* for the GEE and ALR respectively. However, the empirically corrected standard errors for ALR model are somewhat smaller than their counterpart under the GEE model. This implies that the ALR fits the data with small disturbance than GEE. Moreover, ALR extends beyond classical GEE in the sense that precision estimates follow for both the regression parameters **β** and the association parameters **α**. We were also in a position to emphasize that the association is strongly significant ($P < 0.0001$), provided it has been correctly specified, a declaration we could not make in the corresponding exchangeable GEE analysis.

Therefore, we can conclude that ALR is the better model for explaining the marginal association between delivery service and the selected predictor variables. Thus, our interpretation of parameters is based on the final proposed ALR model. Overall, parameter estimates under ALR are slightly less than those of GEE. This difference in parameter estimates from the two models might be due to the fact that ALR takes the associations into account, where as GEE not consider the association parameter in the model.

Table 4.4 Parameter estimates (empirically corrected standard errors) from ALR

Effects	***Level***	***Parame***	***Estimates (s.e)***	***Log 95 % conf.int***	***P-value***
Intercept		$\boldsymbol{\beta_0}$	-3.2778(0.2431)	(-3.7543,-2.8013)	<0.0001
Age	15-19 (Ref)		.	.	.
	20-39	$\boldsymbol{\beta_1}$	0.4978(0.2098)	(0.0865,0.9090)	0.0177
	40-49	$\boldsymbol{\beta_2}$	0.5536(0.2534)	(0.0570,1.0502)	0.0289
Residence	Rural (Ref.)		.	.	.
	Urban	$\boldsymbol{\beta_3}$	2.4599(0.1338)	(2.1976,2.7222)	<.0001
Medul	No education (Ref)		.	.	.
	Primary	$\boldsymbol{\beta_4}$	0.4676(0.0980)	(0.3055,0.6897)	<.0001
	Secondary & above	$\boldsymbol{\beta_5}$	1.3032(0.1643)	(0.9812, 1.6251)	<.0001
Religion	Orthodox (Ref)		.	.	.
	Protestant	$\boldsymbol{\beta_6}$	-0.1482(0.1334)	(-0.4097, 0.1133)	0.2667
	Muslim	$\boldsymbol{\beta_7}$	-0.6312(0.1240)	(-0.8742, -0.3882)	<.0001
	Other	$\boldsymbol{\beta_8}$	-0.3477(0.2610)	(-0.8594, 0.1639)	0.1828
Wealth	Poor (Ref)		.	.	.
	Middle	$\boldsymbol{\beta_9}$	0.3063(0.1228)	(0.0656, 0.5469)	0.0126
	Rich	$\boldsymbol{\beta_{10}}$	0.6666(0.1373)	(0.3973, 0.9357)	<.0001
Brthord	1 (Ref)		.	.	.
	2-4	$\boldsymbol{\beta_{11}}$	-0.6963(0.1091)	(-0.9100, -0.4826)	<.0001
	≥5	$\boldsymbol{\beta_{12}}$	-0.8936(0.1250)	(-1.1386, -0.6487)	<.0001
Hedul	No education (Ref)		.	.	.
	Primary	$\boldsymbol{\beta_{13}}$	0.3063(0.0991)	(0.1120, 0.5005)	<.0001
	Secondary & above	$\boldsymbol{\beta_{14}}$	0.6274(0.1212)	(0.3898, 0.8650)	<.0001
Massme	Not at all (Ref)		.	.	.
	L. than once a week	$\boldsymbol{\beta_{15}}$	0.3776(0.1222)	(0.1380,0.6172)	0.0020
	At least once a week	$\boldsymbol{\beta_{16}}$	0.6893(0.1062)	(0.4812,0.8974)	<.0001
Alpha		$\boldsymbol{\alpha}$	1.0588(0.1024)	(0.8582, 1.2595)	<.0001
		QIC	4123.1463		

4.2.4 Parameter Interpretation of Marginal Models

Table 4.4 presents parameter estimates and their corresponding empirically corrected standard errors alongside the p-values from ALR model. Each parameter $\boldsymbol{\beta_j}$ reflects the effect of factor X_j on the log odds of the probability of getting delivery service, statistically controlling all the other covariates in the model. Then, the odds ratio of variables is calculated as the exponent of $\boldsymbol{\beta_j}$ i.e odds ratio = $\exp(\beta_j)$.

The ALR analysis from table 4.4 suggests that, education is significantly related to delivery service utilization of mothers. After controlling all other variables in the model, the odds ratio of using delivery care service of mothers who has primary education level is $\exp(\beta_4)$= exp(0.4676) = 1.6 (95% CI: 1.357, 1.993) times higher than compared to those non educated mothers, and the odds ratio of secondary and above education level mothers is exp(1.3032) = 3.681 (95% CI: 2.668, 5.1) times higher when compared with non educated mothers (the reference group).

This means that the probability of delivery service of mothers with primary education level is 60% more likely than uneducated mothers and the probability of service delivery is approximately three folds more likely than uneducated mothers.

As we have seen from the result of the ALR model, type of place of residence is statistically significant on delivery service utilization of mothers. The odds ratio of delivery service of mothers living in urban area is exp (2.4599) = 11.7 (95% CI 9, 15.214) times higher than mothers who lives in rural area. This means that the probability of delivery service of mothers who lives in urban area is around 11 times more likely than mothers who live in rural area.

There is also a strong association between age and the use of delivery services utilization of mothers. This implies that, after adjusting all other predictor variables in the model, the estimated odds ratio of using delivery care service for mother's age group 20 to 39 is given as *exp* (0.4978) =1.645 (95% CI: 1.09, 2.482) times and the odds ratio for age group of 40 to 49 is *exp*(0.5563)= 1.74 (95% CI: 1.059, 2.858) times higher when compared with the reference category age 15 to 19. This means that delivery service is increased by 65% and 74% for middle and old age mothers respectively compared to early age group mothers.

Statistically significant difference has been seen by religious view between Orthodox & Muslim mothers. The estimated odds ratio of using assistance during delivery for the Muslim religion mothers, is *exp*(-0.6312) = 0.532 (95% CI: 0.417, 0.678) times lower than Orthodox mothers. This implies that the probability of using delivery service is reduced by 47% for mothers who follow Muslim religion compared with the counter part of Orthodox mothers, keeping the other variables constant in the model. As we have seen that protestant and other religious view are not significant on delivery service use.

Assistance service during delivery showed a decreasing trend with increase in birth order. The odds ratio of using delivery care service utilization is exp (-0.6963) = 0.498 (95% CI: 0.403, 0.617) times lower for birth order of 2 to 4 compared with the first birth order. Similarly the estimated odds ratio of service delivery is exp (-0.8936) = 0.409 (95% CI: 0.320, 0.523) times lower than for birth order five and above as compared with reference group of birth order one. This implies that the delivery service of mothers is reduced by 50% or by half for birth order 2 to 4 and reduced by 59% for birth order five and above compared with first birth order.

Another significant ingredient of mother's delivery care service utilization is wealth index. Mothers from the highest wealth index are more likely to receive delivery care services than those in the poor category. The odds ratio of delivery service of mothers with middle wealth is exp(0.3063) = 1.358 (95% CI: 1.068, 1.728) times higher than mothers with poor wealth index category. Similarly the estimated odds ratio of delivery service of rich mothers, is *exp*(0.6666) = 1.947 (95% CI: 1.488, 2.549) times higher than mothers from poor wealth index. This means that the probability of delivery assistance of mothers from middle wealth index is about 36% more likely than the poor mothers and the delivery service of rich mothers is 95% more likely than the poor counterpart.

Exposure to mass media also another influential predictor variable, for delivery care service of mothers. The odds ratio of delivery service of mothers who follow mass media less than a week, is exp(0.3776) = 1.459 (95% CI: 1.148, 1.854) times higher than mothers who follow mass media not at all. At the same time the odds ratio of delivery service of mothers who expose to mass media at least once a week is exp(0.6893) = 1.992 (95% CI: 1.618, 2.453) times higher compared with mothers not exposed to mass media not at all. This means that the probability of delivery service is 46% more likely for mothers who follow mass media less than once a week and the probability is increased by double for mothers follow mass media at least once a week than who follows mass media not at all controlling the other predictor variables in the model.

Statistically significant association has been seen between partners education level and delivery service utilization of mothers. The odds ratio of delivery service of mothers whose partner has primary education level is exp(0.3063) =1.358, (95% CI: 1.119, 1.650) times higher compared to mothers whose partner has no education. The estimated odds ratio of delivery service of mothers whose partner has secondary and above education level is exp(0.6274)=1.873, (95% CI: 1.477, 2.375), times higher than mothers with whose partner have no any education level. This implies that the probability of delivery service of mothers whose partner has primary education level is 35.8% more likely than with the counter part of mothers with not educated partner and the probability of service delivery of mothers whose partner has secondary and above education is 87% more likely than mothers with uneducated partner.

The ALR model also presents the estimated constant log odds ratio (alpha) which, provide information about the association between individual observations within the same cluster. The estimated pair wise odds ratio relating two responses from the same cluster is exp(1.0588) =2.883 (95% CI: 2.360, 3.524). These, the value of alpha which is greater than one indicates that, the associations is found to be significant (p-value <.0001) and this means that there is a strong positive association between individual mothers about delivery service use in the same cluster.

4.3 ANALYSIS OF GENERALIZED LINEAR MIXED MODEL (GLMM)

4.3.1 Model Building in GLMM

Under the GLMM, model fitting began by adoption of the marginal model covariates. Additionally, the model also included the random effects in this case, random intercepts to address the between and within-regional variations. First, all main effect covariates and the two random intercepts model were fitted and as usual, non significant covariates were removed sequentially starting from variables with highest p-value for fixed effect covariates. Then the saturated models for GLMM were fitted as follows where, v_j & u_{ij} two random intercepts.

$$\begin{aligned} logit(\pi_{ij}) = {} & \beta_0 + \beta_1 Age_1 + \beta_2 Age_2 + \beta_3 Residence_U + \beta_4 Medul_P + \beta_5 Medul_{S+} \\ & + \beta_6 Religion_{Pr} + \beta_7 Religion_{Mu} + \beta_8 Religion_O + \beta_9 Sexhh_M + \beta_{10} Wealth_M \\ & + \beta_{11} Wealth_R + \beta_{12} Brtord_{2+} + \beta_{13} Brtord_{5+} + \beta_{14} Hedul_P + \beta_{15} Hedul_{S+} \\ & + \beta_{16} Workst_E + \beta_{17} Massme_L + \beta_{18} Massme_{At} + \boldsymbol{v_j} + \boldsymbol{u_{ij}} \end{aligned}$$

In order to decide on the better of the two random effects models, two models were fitted, one the saturated model above with two random intercepts to estimate between and within regional variations and the other with one random intercept model to estimate within regional variation. AIC and Likelihood ratio test (LRT) were used to compared the two models to select an appropriate models.

Table 4.5 Information criteria for comparison of one and two random intercept models

Models	**AIC**	**BIC**	**-2LogLik**	**Deviance**	$\boldsymbol{\sigma_W}$	$\boldsymbol{\sigma_B}$	**P**
One random intercept model	3854	3993	3814	3814	1.1578		
Two random intercept model	3797	3943	3754	3755	1.0086	0 .7804	0.000

Where, σ_W and σ_B are within and between regional standard deviation respectively, and P is the p-value of the log likelihood ratio test of the two models.

As we have seen from table 4.5, the AIC of two random intercept model is reduced from 3854 to 3797, the -2loglikelihood is reduced from 3814 to 3754 & the deviance of the model is reduced from 3814 to 3755. The small p-value of the log likelihood ratio test ($P < 0.001$) also indicates that the model with two random intercept is parsimonious model.

Also when considered a model without random effects (i.e simply the generalized linear model), it gives AIC value of 4053.833 which is large as compared to the above two models with random effects. In addition, the likelihood ratio test at the bottom panel of table 4.6 in GLMM parameter estimate output also shows that the comparison of random effect model versus the ordinary logistic model (GLM) without random effects. The resulting p-value (P < 0001) of this test supports that considering the random effect model is essential. Therefore, we conclude that, the model with two random intercepts should be used to address the between and within-regional heterogeneity in the given data.

Next, the covariates for the fixed effect were assessed and the candidate covariates were selected by removing covariates starting from with highest p-value sequentially. Then the first removable covariate is work status of mothers with the highest p-value (P = 0.8799) and refitted the reduced model with the remaining covariates. The AIC is reduced from 3795 to 3793 and the p-value of log likelihood ratio test (P = 0.8811) supports the reduced model is preferable one. The next removable variable is sex of the household leader with p-value (P = 0.1181) and refitted the reduced model. For this model, AIC is similar with the previous one but the likelihood ratio test indicates that the reduced model is better with the p-value (P=0.1228). In addition, the model with small number of covariates is considered to be preferable. Therefore, the final proposed GLMM for delivery care service utilization of mothers is given as:

$$\begin{aligned} logit(\pi_{ij}) = \beta_0 + \beta_1 Age_1 + \beta_2 Age_2 + \beta_3 Residence_U + \beta_4 Medul_P + \beta_5 Medul_{S+} \\ + \beta_6 Religion_{Pr} + \beta_7 Religion_{Mu} + \beta_8 Religion_O + \beta_9 Wealth_M \\ + \beta_{10} Wealth_R + \beta_{11} Brtord_{2+} + \beta_{12} Brtord_{5+} + \beta_{13} Hedul_P + \beta_{14} Hedul_{S+} \\ + \beta_{15} Massme_L + \beta_{16} Massme_{At} + \boldsymbol{v_j} + \boldsymbol{u_{ij}} \end{aligned}$$

The parameter estimates and standard errors of the GLMM are presented in table 4.6.

4.3.2 Parameter Interpretation of GLMM

Unlike in the marginal models, (GEE and ALR) where parameters are treated as population averages, in the GLMM analysis, parameter interpretation is based on specific subjects or cluster. The parameter interpretation is conditional on the random effects, which is common for all individual mothers in the same cluster.

Table 4.6 Parameter estimates (standard errors) and corresponding P value for GLMM.

Effects	**Level**	***Parameter***	**Estimates (S.e)**	***P-value***	**95% conf.int**
Intercept		β_0	*-3.4771(0.3515)*	*0.000*	*(-4.1661,-2.7882)*
Age	15-19 (Ref)		.	.	
	20-39	β_1	*0.6148(0.2092)*	*0.003*	*(0.2048, 1.0249)*
	40-49	β_2	*0.6864(0.2940)*	*0.020*	*(0.1102, 1.2626)*
Residence	Rural (Ref.)		.	.	
	Urban	β_3	*2.3279(0.1703)*	*0.000*	*(1.9942, 2.6616)*
Medul	No education (Ref)		.	.	
	Primary	β_4	*0.5872(0.1124)*	*0.000*	*(0.3638, 0.8076)*
	Secondary & above	β_5	*1.6056(0.1975)*	*0.000*	*(1.2186, 1.9926)*
Religion	Orthodox (Ref)		.	.	
	Protestant	β_6	*-0.3387(0.1703)*	*0.044*	*(-0.6724, -0.0049)*
	Muslim	β_7	*-0.7662(0.1539)*	*0.000*	*(-1.0679, -0.4645)*
	Others	β_8	*-0.5449(0.3378)*	*0.107*	*(-1.2070, 0.1171)*
Wealth	Poor (Ref)		.	.	
	Middle	β_9	*0.3330(0.1490)*	*0.025*	*(0.0410, 0.6249)*
	Rich	β_{10}	*0.4995(0.1679)*	*0.003*	*(0.1705, 0.8285)*
Brthord	1 (Ref)		.	.	
	2-4	β_{11}	*-0.8078(0.1196)*	*0.000*	*(-1.0423, -0.5733)*
	≥5	β_{12}	*-1.012(0.1475)*	*0.000*	*(-1.3007, -0.7225)*
Hedul	No education (Ref)		.	.	
	Primary	β_{13}	*0.3446(0.1165)*	*0.003*	*(0.1163, 0.5729)*
	Secondary & above	β_{14}	*0.7257(0.1545)*	*0.000*	*(0.4228, 1.0286)*
Massme	Not at all (Ref)		.	.	
	Less than once week	β_{15}	*0.4313(0.1405)*	*0.002*	*(0.1558, 0.7069)*
	At least once a week	β_{16}	*0.8693(0.1287)*	*0.000*	*(0.6170, 1.1216)*
Random effects		σ_W	*1.013(0.0798)*		*(0.8681, 1.2516)*
		σ_B	*0.7812(0.1879)*		*(0.4876, 1.2184)*
LR test vs. logistic regression:			*Chi2(2)=266.51*		*Prob > chi2 = 0.0000*

Ref=reference category

Given the same random effects b_j, the estimated odds ratio of delivery service use of mothers is exp (0.6148) = 1.849 (95% CI: 1.227, 2.787) times higher for age group 20-39 and exp (0.6864) = 1.987 (95% CI: 1.117, 3.535) times higher for age group 40-49 compared to mothers with age group 15-19 in the same j^{th} cluster keeping constant the other fixed effect variables in the model. This implies that the probability of delivery service use is 85% and 98.7% more likely for mothers whose age group is 20-39 & 40-49 respectively than with mothers whose age group is 15-19 in the same cluster at the given random effects.

In the same way, the estimated odds ratio of delivery service of mothers was exp (-0.3387) = 0.713 (95% CI: 0.510, 0.995) and exp(-0.7662) = 0.465 (95% CI: 0.344, 0.628) times lower for Protestant and Muslim religious mothers respectively compared with Orthodox religion mothers in the same j^{th} cluster with constant random effect in the given cluster and the other fixed effect covariates in the model are constant.

At the given constant random effect, the odds ratio of delivery care service utilization of mothers is exp(-0.8078) = 0.446 (95% CI: 0.353, 0.564) times lower for the birth order group of 2-4 and exp (-1.012) = 0.363 (95% CI: 0.272, 0.456) times lower for birth order five and greater than mothers with first birth order in the same cluster. This shows that the probability of delivery service for mothers with the birth order group 2 to 4, and birth order five and greater, is 55.4% and 64.7% less likely than mothers with first birth order respectively at the same cluster with the same random effect. Except the variable place of residence, the interpretation of other predictor variables can be done in a similar manner.

Since clustering for 2011 EDHS was considered urban and rural area, parameter interpretation of the covariate, type of place of residence is at regional level random effects. Then, the odds ratio of delivery care service of mothers who lives in urban place is exp (2.3279) = 10.256 (95% CI: 7.346, 14.319) times higher than mothers who lives in rural area in the same region keeping constant other covariates and regional level random effects. This implies that the probability of delivery service for urban area is around nine folds more likely than rural mothers in the given region.

The random effect parameters under GLMM are not estimable and then we cannot interpret it. However the estimates of within and between standard deviation of random effects are 1.013 and 0.7812 which is larger than zero & the boundary of 95% confidence interval for the estimates doesn't close to zero. In addition, the likelihood ratio test of GLMM versus ordinary logit is highly significant ($P < 0.0001$). Then, we can interpret; there is significance heterogeneity within and between regions on the delivery service use of mothers.

4.3.3 Model diagnostic for GLMM

The Q-Q plot from the following figure in first panel verifies that the residuals are close to normally distributed and symmetric around zero. Thus, it meets the assumption of the distribution of error terms. As well, to the above, the non linearity of the Q-Q plot confirms the model is not linear. Residuals versus observation CLID number plot panel two, also suggested that the residuals are symmetric around zero (i.e. positive and negative residuals are almost equal). Q-Q plots for normality of random effects at regional and cluster levels are also given in the figure at panel three and four, and illustrates that the random effects are normally distributed with mean zero and variance covariance matrix D. Thus, the fitted GLMM model is fine for the given data.

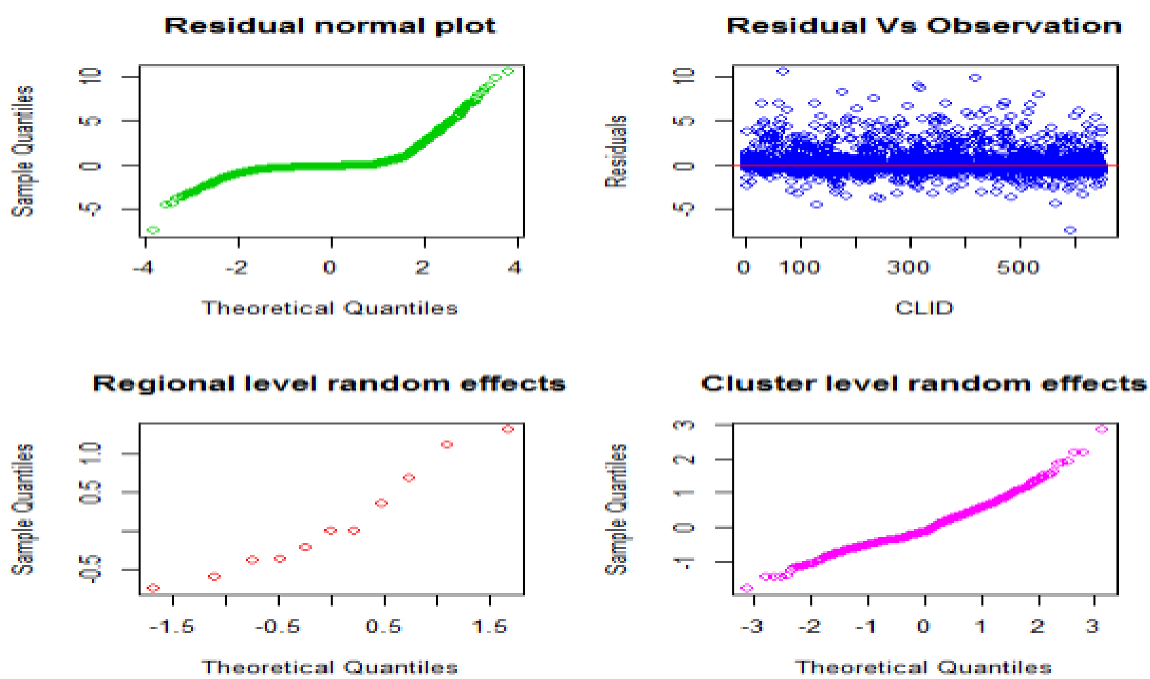

Figure 4.1 Diagnosis plots for the generalized linear mixed model

4.4 DISCUSSION

This study was aimed at modeling the delivery care service utilization of mothers in Ethiopia. As a preliminary analysis, assortments of summary statistics were employed to explore the association between the response variable of interest and available covariates. It should be well-known that there is inconsistency in the conclusion from the analysis of various summary statistics, which might be due to the fact that they make use of varying amount of information, which determines the power of their inferences. Thus, the analysis was extended to other statistical methods to account for the clustered nature of correlated observations. The data were then analyzed using two model families one with marginal models (GEE and ALR), and the other is random effects model (Generalized linear mixed model).

Two proposed working correlation structures, exchangeable and independence correlation assumptions were taken for the comparison, in GEE model building strategy. The model with exchangeable working correlation structure was found to be better fits the data than independence. This supports that considered the clustering nature of the data was essential for the analysis and the dependency of individuals for the given data. In addition, ALR was fitted for simultaneously regress the response variable on explanatory variables as well as association among responses in terms of pairwise odds ratio.

Two models from marginal model families were compared in order to assess which model is efficiently explain the relations between response and explanatory variables as well as to evaluate that whether considering pairwise association is important. After then, ALR model was selected as best model and the model shows that there is a positive pairwise association between responses. This is supported the idea explained by Zeger *et al*, alternating logistic regression is reasonably efficient relative to GEE (Zeger *et al*, 1993). Overall, parameter estimates and its standard errors under ALR are slightly less than those of GEE. These differences in parameter estimates from the two models might be due to the fact that ALR takes the associations into account, where as GEE not consider the association parameter in the model, which supports findings of Cosmas (2011).

The purpose of GLMM was to evaluate within and between regional variations of delivery service utilization of mothers in Ethiopia. Two models was fitted one with only one intercept model to assess only within regional variation and other with two random intercepts model, in order to account within and between regional variations. Additionally, generalized linear model was fitted, as the sake of comparison whether including random effects in the analysis is important. The three models were compared using the AIC value followed by likelihood ratio test and we got a model with two random intercept was favorable. This demonstrates that, accounting within and between regional variations for the analysis of delivery care services of mothers should be vital and, indicates within and between regional heterogeneity in delivery care service utilization of mothers. This finding is supported by the explanation or suggestion of Antonio & Beirlant (2006).

However, that the two model families are different and their comparability may not be meaningful as they have different parameter interpretations and estimations, parameter estimates obtained from GLMM are generally bigger in absolute values than those from marginal models (GEE and ALR) similar with Agresti (2007).

All the fitted models were leads to the same conclusion that age, residence, mother's education level, religion, wealth index, birth orders, partner's education level and exposure to mass media were found to be significantly associated with delivery service utilization of mothers.

Age has a positive association with the use of delivery assistance. Older women are more likely to use skilled delivery assistance as compared to younger women. This can be explained by the fact that, older women are generally more experienced and knowledgeable about healthcare services and their use, which may improve utilization. Older women may also be more confident and have higher household decision-making power than that of the younger women, particularly adolescents, which will improve their likelihood of health service use. This is similar with studies by Reynolds *et al*, (2006) and Abor *et al*, (2011).

Similar with the previous studies Elo (1992), Kistiana (2009), Umurungi (2010), this study confirmed that a significantly positive association between education of mothers and the use of skilled assistance during delivery. Education serves as prospect for information and knowledge of available health care services utilization. Education also serves as proxy for women's higher socioeconomic status that improves the ability of educated mothers to afford the cost of health care services utilization. Moreover, educated women are considered to have better knowledge and information on modern health care services. These factors, therefore, enable women to seek for safer childbirth under the supervision of skilled attendants. Women's education was found to be a strong determinant of the use of skilled assistance at delivery.

Similar with women's education, we also found husband's education as an important predictor of the use of assistance during delivery. This finding conforms to some previous studies in Ethiopia, Eyerusalem (2010) and in Indonesia, Kistiana (2009). It is likely that an educated family will have a better understanding and knowledge of modern health care services. Education also leads to better awareness of available services. These, in turn, sensitize the educated family to make use of available services including maternal health services whenever they perceive it to be necessary (Abul 2012).

As several studies, we also found that negative association between birth order and the use of skilled delivery assistance during delivery. This is consistence with findings from Mesfin et al (2004), Ethiopian Society of Population Studies (ESPS, 2008) in Ethiopia and Wang et al (2011) in sub Saharan Africa, delivery service decreases with higher birth order. This can be explained by a reason that fear of complication or lack of confidence is of mothers who experience first birth and thus, are more likely to use delivery assistance at the time of delivery than mothers with higher birth order. Conversely, mothers with more children believe themselves to be more experienced in childbirth, hence, are less likely to use skilled assistance at delivery. The low use of skilled birth assistance at delivery among mothers of higher number of children can also be due to the resource limitation in the family as there are many demands in the family (Abul 2012).

In this study, urban mothers were significantly associated with increased odds of delivering with skilled assistance. This finding reflects the finding of several previous studies which have reported a significantly higher use of skilled assistance at delivery by urban mothers compared to rural mothers in Ethiopia, (Mekonen 2002, Eyerusalem, 2010 and Asmeret 2013) and Kistiana (2009) in Indonesia. A reason for this may be the availability of health facilities, because health facilities are much more convenient in urban areas than rural areas in developing countries like Ethiopia. This close proximity allows urban mothers greater access to information and knowledge regarding modern health care facilities, which influences them to use these facilities. Other reasons may be that the urban mothers are from the families who have a higher level of education and have a higher level of household economic status (Abul 2012).

This study finding regarding the positive association between wealth quintile of mothers and the use of delivery assistance during delivery coincides with previous studies in Ethiopia Ethiopian Society of Population Studies (2008) and Asmeret (2013), in Rwanda Umurungi (2010) and in Ghana Charles *et al* (2011). We found that the odds of using delivery service at the time of delivery consistently increased as the household economic status increased. A reason for this finding may be that the family members from higher level of household economic status are more aware of accessible modern health care services and can afford those services easily. The costs of seeking skilled assistance at delivery may act as an important barrier to mothers from poor households.

This study shows that, mothers exposure to mass media were more likely to use delivery services from health professionals at delivery than mothers who doesn't expose to mass media. The results confirm similar findings from a study in Indonesia Kistiana (2009) and south India Navaneetham & Dharmalingam (2000). This might be due to the fact that, the role of mass media in changing both patterns of delivery service use and planning of ideal health care programs among those mothers exposed to mass media. Mothers with information about maternal health care programs through radio, on TV or in news papers can create awareness and improving knowledge about the importance of delivery care service.

We found that mothers who follow Muslim religious are less likely to use skilled assistance at delivery compared to their counterpart of Orthodox Christian mothers. The finding conform to study by Anwar et al indicating that Muslim mothers are less likely to use skilled assistance at delivery compared to mothers from other religions Anwar *et al* (2008). A possible reason for this finding may be the local tradition and culture that influence Muslim mothers not to use delivery assistance during delivery.
However, from the previous studies, work status of mothers and sex of house head were significantly associated with delivery service utilizations, these covariates doesn't significant determinant factors on this study.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

This study analysis the association between delivery care services of mothers and its possible determinant factors as well as tried to answer important issues on the status of mother's delivery service utilization in Ethiopia by applying reasonably applicable statistical models. Mothers' delivery service utilization was found to be very low and it was seen that use of this services were unequally distributed. A total of 7737 eligible mothers from EDHS 2011 data were included in the study. Among these, the proportion of coverage of delivery service is only 17.2%, which is extremely near to the ground.

Based on the clustering nature of the data, two model families, marginal (GEE & ALR) and cluster specific (GLMM) models were applicable for the appropriate analysis in this study. The primary scientific objective of marginal models is to analyze the population-averaged effects of the given factors in the study on the binary response variable of interest for clustered data. In addition, the purpose of ALR model was to measures that provided information about the association between individual observations within the same cluster. For this study two marginal models, GEE and ALR, have been compared for the analysis of marginal or average effects of covariates on the response variable and, we conclude that, ALR model with measure of association exhibited the best fit for this data than GEE models.

Cluster specific (GLMM) model was applied in this study for the purpose of including random effect parameters specific to clusters, which are directly used in modeling the random variation in the dependent variable at different levels of the data. For this study GLMM, with two random intercept model was found to be appropriate for the analysis of within and between regional variations for delivery service utilization of mothers in Ethiopia. This concluded that there is heterogeneity by delivery service utilization of mothers between and within regions.

Age has a positive association with delivery service utilization of mothers that the probability of delivery service of mothers is more likely for elder age group than that of adolescent age.

In this study, analysis indicated that parental education as the most significant predictive factors for the use of delivery assistance at the time of delivery. The probability of delivery service increases with the higher level of education of mothers and partners. There is a negative association between birth order and delivery assistance of mothers. Higher birth order appeared as a strong predictor to not use of professional assistance during delivery.

The study concludes that mothers from higher economic status have greater probability to use delivery assistance at the time of delivery. We also conclude that place of residence creates a great barrier on the delivery service utilization of rural mothers in Ethiopia. The probability of delivery assistance of urban mothers is more than 9 folds compared with mother who lives in rural Ethiopia. In addition, we conclude that mass media has a very significant impact on the delivery care services utilization; mothers who exposed to mass media have more probability for delivery service than not exposed to mass media.

5.2 RECOMMENDATION

Since delivery service utilization of mothers, is a critical and current issue to reduce the risk of complications and infections that can cause the death or serious illness of the mother and the newborn baby, these modeling this service utilization have an important policy implications. Most of the researchers often interested on ordinary logistic regression model in the field of medical and other sciences. However, it always does not satisfy the assumptions by the nature of the data and can be leads to unreliable conclusion. Therefore, it should be consider the nature of the data and applied appropriate statistical model families, which gives relevant outputs and statistical inferences like marginal and generalized linear mixed model for clustered data.

This study has identified a number of important factors that influence the use of assistance during delivery of mothers in Ethiopia. Parental education is one of the most significant predictive factors for the use of delivery assistance of mothers. Therefore, informal adult education for mothers and partners should be employed as an immediate intervention to provide basic education and to increase awareness about basic maternity health care. Besides, special efforts and attention to improve formal education of the girls and boys are needed in a long run.

Higher birth order appeared as a strong negative predictor to use of delivery assistance of mothers. Therefore, raising awareness about the use of delivery service among mothers and partners through mass media and local human resources (religious leader, political leader, school teacher, village headman, and singer) should be an immediate intervention accompanied by improving access to family planning as a long term strategy. Mothers from rural areas and mothers from lower wealth index were at a greater disadvantage in using delivery assistance. Therefore, informal education and vocational training for those groups of mothers may serve as an immediate strategy to improve the use of delivery service utilization. Moreover, special attempts of delivery services should be prepared for those groups of mothers as a long term strategy. More services should be offered to the rural areas with mass awareness program to use those services.

REFERENCE

Abor, A., Abekah-Nkrumah, G., Saky, K., Adjasi, D., & Abor, J., (2011).The socio-economic determinants of maternal health care utilization in Ghana, *International Journal of Social Economics, **V**ol. 38,No. 7, pp. 628-648.*

Abul, B., (2012).Department of Public Health and Clinical Medicine Epidemiology and Global Health, Umea International School of Public Health, Umea University, Sweden

Adamu, S., (2011).Utilization of Maternal Health Care Services in Nigeria:, An Analysis of Regional Differences in the Patterns and Determinants of Maternal Health Care Use, University of Liverpool.

Agresti, A., (2007).An Introduction to Categorical Data Analysis, 2nd Ed, Wiley Inc

Agresti, A., (2002).Categorical Data Analysis, Second Edition

Anna, J., (2007).Why are there low institutional delivery rates in the Gambia? Women's opinion, Institute of General Practice and Community Medicine, Section for International Health, University of Oslo

Anderson, J., Verkuilen, J., Johnson, R., (2012).Applied Generalized Linear Mixed Models:, Continuous and Discrete Data For the Social and Behavioral Sciences.

Anwar I, Sami M, Akhtar N, Chowdhury ME, Salma U, Rahmanc M, Koblinskya M., (2008). Inequity in maternal health-care services: evidence from home-based skilled-birth attendant programs in Bangladesh. Bulletin of the World Health Organization; 86 (4):252-259.

Antonio, K. & Beirlant, J., (2006).Actuarial Statistics with Generalized Linear Mixed Models, University Center for Statistics, Belgium.

Asmeret, M., (2013).Levels and Determinants of Use of Institutional Delivery Care Services among Women of Childbearing Age in Ethiopia:, Analysis of EDHS 2000 and 2005 Data ICF International Calverton, Maryland, USA.

Babalola, S. & Fatusi, A., (2009).Determinants of use of maternal health services in Nigeria, looking beyond individuals and household factors

Baral, Y., Lyons, K., Skinner, J., & Teijlingen R., (2010).Determinants of skilled birth attendants for delivery in Nepal, ***Vol.8***.

Carey, V., Zeger, L., Diggle, P., (1993).Modeling Multivariate Binary Data with Alternating Logistic Regressions, *Biometrika, **V**ol. 80, pp. 517-526*.

Central Statistical Agency (CSA), and ORC Macro, (2012), Ethiopia Demographic and Health Survey 2011, Addis Ababa, Ethiopia, and Calverton, Maryland, USA:, Central Statistical Agency and ORC Macro.

Cosmas, G., (2011).Socio-Demographic determinants of anemia among children aged 6-59 months in mainland Tanzania, University of Hasselt.

Elo, T., (1992).Utilization of maternal health-care services in Peru:, The role of women's education, *Health Transition Review,* **V**ol.*2, pp. 49–69.*

Ethiopian Society of Population Studies, (2008).Maternal Healthcare Seeking Behavior in Ethiopia:, Findings from EDHS 2005, Addis Ababa.

Eyerusalem, D., (2010).Role of Socio-Demographic Factors on Utilization of Maternal Health Care Services in Ethiopia, Umea University.

Faraway, J., (2006).Extending the Linear Model with R, Generalized Linear, Mixed Effects, and Nonparametric Regression Models, *Boca Raton London New York,*

Garg, R., Shyamsunder, D., Singh, T., Singh, A.,(2010).Study on Delivery Practices among Women in Rural Punjab, *Health and Population **V**ol.33,pp. 23-33*.

Idris, S., Gwarzo, D., & Shehu, U.,(2006).Determinants of Place of Delivery among Women in a Semi-Urban Settlement in Zaria Northern Nigeria, *Annuals of African Medicine pp. 68-72.*

JSI Research & Training Institute JSI, (2009). Maternal Health Seeking Behaviors in Ethiopia:, Findings from L10K Baseline Survey, Arlington, VA, USA: JSI.

Kassu, M., (2012).Determinant Factors Affecting Utilization of Maternal Health Care Services in Rural Ethiopia, Addis Ababa University, Ethiopia

Kesterton, A., Cleland, J., Sloggett, A., &Ronsmans, C., (2010).Institutional Delivery in Rural India: The Relative Importance of Accessibility and Economic Status, BMC Pregnancy & Child Birth.

Kistiana, S., (2009).Socio-economic and demographic determinants of maternal health care utilization in Indonesia.

Macro International, (2007).Trends in Demographic and Reproductive Health Indicators in Ethiopia, Calverton, MD, USA: Macro International Inc.

McCullagh, P., & Nelder, J. (1989).Generalized Linear Models, 2nd edition, London, Chapman, and Hall

McCulloch, E., (1997).An Introduction to Generalized Linear Mixed Models, Biometrics Unit, and Statistics Center Cornell University.

Mekonnen, Y., &Mekonnen, A., (2002).Utilization of Maternal Health Care Services in Ethiopia, Calverton, MD, USA: ORC Macro.

Mesfin, N., Damen, M., & Getnet, M.,(2004).Assessment of Safe Delivery Service Utilization among Women of Childbearing Age in North Gondar Zone, North West Ethiopia, *Ethiop J Health Dev18(3): 146-52*.

Molenberghs, G., & Verbeke, G., (2005).Models for Discrete Longitudinal Data. Library of Congress

Montagu, D., Yamey, G., Visconti, A., Harding, A., & Yoong, J., (2011). Where Do Poor Women in Developing Countries Give Birth? A Multi-Country Analysis of Demographic, and Health Survey Data

Myers, H., Montgomery, C., Vining, G. &Robinson, J., (2010).Generalized Linear Models, With Applications in Engineering and the Sciences, Second Edition.

Navaneetham, K., & Dharmalingam, A., (2000). Utilization of Maternal Health Care Services in South India, Paper presented in faculty seminar at the Centre for Development Studies, Thiruvananthapuram, Kerala.

Neuhaus, J., Kalbfleisch, J. & Hauck, W., (1991).A Comparison of Cluster-Specific and Population-Averaged Approaches for Analyzing Correlated Binary Data, *International Statistical Review, Vol. 59, pp. 25-35.*

Pan, W., (2001).Akaike's information criterion in generalized estimating equations, Biometrics, 57, 120-125

Stephenson, R., Baschieri, A., Clements, S., Hennink, M., & Madise, N., (2006). Contextual Influences on the Use of Health Facilities for Child Birth in Africa, *American Journal of Public Health, pp., 84-93.*

Susan, O. & Elizabeth, R., (2010).Bringing Care Closer to Mothers and Newborns: *Using the Gap Analysis Tool to Develop a Household to Hospital Continuum of Care, save the children*

Thind, A., Mohani A., Banerjee K., Hagigi F.,(2008*)*.Where to deliver? Analysis of choice of delivery location from a national survey in India, *Bio Med Central Public Health*

Umurungi, S., (2010).Determinants of the utilization of delivery services by pregnant women in Rwanda.

Wang, W., Alva, S., Wang, S. & Fort, A., (2011).Levels and Trends in the Use of Maternal Health Services in Developing Countries, *DHS Comparative Reports No. 26, ICF Macro Calverton, Maryland, USA.*

World Health Organization (WHO), UNICEF, UNFPA, & World Bank, (2008).Trends in Maternal Mortality: 1990-2008: Estimates Developed by WHO, UNICEF, UNFPA and the World Bank, Geneva, Switzerland: WHO.

Wedderburn, R.W.M. (1974).Quasi-Likelihood Functions, Generalized Linear Models and the Gauss Newton Method," Biometrika, 61, pp.439-447.

Zeger, S. & Liang K., (1986). Longitudinal data analysis using generalized linear models. *Biometrics, **Vol.73**, pp.13-22*

Ziegler, A. Kastner, C. Blettner, M. (1998).The Generalized Estimating Equations: An Annotated Bibliography. *Biometrical Journal 40, pp.115-139*

APPENDIX

SAS and R cods for the models

SAS code

```
/* Fitting GEE model*/
proc genmod data=Thesis descending;
class  CLID Region Age(ref="15-19") Residence(ref="Rural") Medul(ref="No
education") Religion(ref="Orthodox")Sexhh(ref="Female") Wealth(ref="Poor")
Brthord(ref="1") Hedul(ref="No education") Workst(ref="Non Employement")
Massme(ref="Not at all") / param=ref ref=first;
model Delivery=Age Residence Medul Religion Wealth Brthord Hedul Massme /
dist=bin link=logit;
repeated subject=CLID /type=exch modelse;
run;
/*Fitting ALR model: */
proc genmod data=Tthesis descending;
class  CLID Region Age(ref="15-19") Residence(ref="Rural") Medul(ref="No
education") Religion(ref="Orthodox")Sexhh(ref="Female") Wealth(ref="Poor")
Brthord(ref="1") Hedul(ref="No education") Workst(ref="Non Employement")
Massme(ref="Not at all") / param=ref ref=first;
model Delivery=Age Residence Medul Religion Wealth Brthord Hedul Massme /
dist=bin link=logit;
repeated subject=CLID /logor=exch modelse;run;
```

R-code for GLMM model

```
library(MASS)
library(foreign)
library(lme4)
deliv<-read.spss("C:\\Aresearch\\Delivery.sav")
delcare<-as.data.frame(deliv)
head(delcare)
attach(deliv)
fit1<-lmer(Delivery~factor(Age)+factor(Residence)+factor(Medul)+
factor(Religion)+factor(Sexhh)+factor(Wealth)+factor(Brthord)+factor(Hedul)+facto
r(Workst)+factor(Massme)+(1|Region)+(1|CLID),family=binomial,data=delcare)
print(fit1,corr=FALSE)
fit2<-lmer(Delivery~factor(Age)+factor(Residence)+factor(Medul)+
factor(Religion)+factor(Sexhh)+factor(Wealth)+factor(Brthord)+factor(Hedul)+facto
r(Workst)+factor(Massme)+(1|CLID),family=binomial,data=delcare)
print(fit2,corr=FALSE)
### Comparison of random effects
anova(fit2,fit1)
### Variable selection
fit3<-lmer(Delivery~factor(Age)+factor(Residence)+factor(Medul)+factor(Religion)+
factor(Sexhh)+factor(Wealth)+factor(Brthord)+factor(Hedul)+factor(Massme)+(1|Regi
on)+(1|CLID),family=binomial,data=delcare)
print(fit3,corr=FALSE)
fit4<-
lmer(Delivery~factor(Age)+factor(Residence)+factor(Medul)++factor(Religion)+
factor(Wealth)+factor(Brthord)+factor(Hedul)+factor(Massme)+(1|Region)+(1|CLID),
```

```
family=binomial,data=delcare)
### Comparison of all models for variable selection
anova(fit1,4fit3,fit4)

##### Diagnosis Plots######
par(mfrow=c(2,2))
qqnorm(resid(fit4),main="Residual normal plot",col=3)
plot(delcare$CLID,resid(fit4),xlab="CLID",ylab="Residuals",main="Residual Vs
Observation",col=4)
abline(h=0,col=2)
qqnorm(ranef(fit4)$"Region"[[1]],main="Regional level random effects",col=2)
qqnorm(ranef(fit4)$"CLID"[[1]],main="Cluster level random effects",col=6)

###STATA 11 ###
xtmelogit in STATA 11 is also used for GLMM analysis
```

The full model Wald test for variable selection in GEE

Score Statistics For Type 3 GEE Analysis

Source	DF	Chi-Square	Pr > ChiSq
AGE	2	6.65	0.0360
RESIDENCE	1	183.81	<.0001
MEDUL	2	54.22	<.0001
RELIGION	3	24.89	<.0001
SEXHH	1	3.44	0.0638
WEALTH	2	29.15	<.0001
BRTHORD	2	46.40	<.0001
HEDUL	2	25.03	<.0001
WORKST	1	0.00	0.9714
MASSME	2	35.68	<.0001

QIC=4100.0609

The full model Wald test for variable selection in ALR

Score Statistics For Type 3 GEE Analysis

Source	DF	Chi-Square	Pr > ChiSq
AGE	2	6.02	0.0493
RESIDENCE	1	185.14	<.0001
MEDUL	2	49.71	<.0001
RELIGION	3	22.88	<.0001
SEXHH	1	1.73	0.1878
WEALTH	2	18.33	0.0001
BRTHORD	2	40.36	<.0001
HEDUL	2	22.30	<.0001
WORKST	1	0.14	0.7047
MASSME	2	33.83	<.0001

QIC=4125.6687