



**Socio-economic and Demographic Determinants of
Unemployment in Ethiopia**

ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

Berhan Abera

A Thesis Submitted to the Department of Statistics

Presented in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Statistics

Addis Ababa, Ethiopia, January 2013

Addis Ababa University

School of graduate studies

This is to certify that the thesis prepared by Berhan Abera, entitled: **Socio-economic and Demographic Determinants of Unemployment in Ethiopia** and submitted in partial fulfillment of the requirements for the Degree master of science in Statistics compiles with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the examining committee:

Examiner_____signature_____Date_____

Examiner_____signature_____Date_____

Advisor_____signature _____Date_____

Chair of Department or Graduate Program Coordinator

ABSTRACT

Socio-economic and Demographic Determinants of Unemployment in Ethiopia

Berhan Abera

Addis Ababa University, 2013

Unemployment is the most attention-seeking challenge that faces the Ethiopian economy of today. The objective of the study is to identify the socio-economic and demographic determinants of unemployment in Ethiopia. Descriptive and binary logistic regression analyses were used for analyzing the data. The predictors age, sex, region, place of residence, educational level, economic status, marital status, sex of head of household and household size are found to have a significance effect on unemployment status of an individual in Ethiopia. It is recommended that, in the short term creating jobs that require manual work by humans is possible. Empower women and increase their participation. Efforts made by the government and other organizations to do so should further be enhanced. In the long run, increasing and improving the level of education will create better chances not to live in poverty.

Acknowledgement

First and foremost I would like to thank God for everything. Next I am grateful to M.K. Sharma (Professor), my thesis advisor, for his support and comments. I am also thankful for Addis Ababa University for giving me the scholarship and the department of Statistics. Finally my gratitude goes to my family and friends.

Acronyms

Cons - Constant

CSA - Central Statistics Agency

EAs -Enumeration Areas

EDHS -Ethiopian Demographic and Health Survey

GDP -Gross Domestic Product

ILO - International Labor Organization

LB – Lower Bound

LFS - Labor Force Survey

LL - Likelihood

LR - Likelihood Ratio

ML - Maximum Likelihood

MoFED - Ministry of Finance and Economic Development

MOH - Ministry of Health

OLS - Ordinary Least Squares

OR - Odds Ratio

SNNP - Southern Nations, Nationalities and People's Region

SPSS -Statistical Package for Social Science

Std.Err - Standard Error

UB – Upper Bound

USD - United States Dollar

WB - World Bank

Table of Contents

ABSTRACT.....	iii
Acknowledgement	iv
Acronyms.....	v
List of Tables	viii
CHAPTER ONE	1
1. Introduction.....	1
1.1 Background of the Study	1
1.2 Statement of the Problem.....	4
1.3 Objective of the Study	5
1.4 Significance of the Study	6
1.5 Limitation of the Study	6
CHAPTER TWO	7
2. Literature Review.....	7
2.1 Unemployment as a Global Problem	7
2.2 Unemployment in Ethiopia.....	11
CHAPTER THREE	13
3. Data and Methodology.....	13
3.1 Data Source.....	13
3.2 Description of Variables	13
3.2.1 The Dependent Variable	14
3.2.2 Independent Variables	14

3.3 Methodology	16
3.3.1 Logistic Regression.....	16
3.3.1.1 Assumptions of the Binary Logistic Regression.....	17
3.3.1.2 The Model.....	18
3.3.1.3 Estimation of Model Parameters.....	20
3.3.1.4 Goodness of Fit of the Model	21
3.3.1.5 Test of Significance of an Individual Predictor	24
3.3.1.6 Model Diagnostic.....	25
CHAPTER FOUR.....	27
4. Statistical Data Analyses.....	27
4.1 Descriptive Data Analysis.....	27
4.2 Logistic Regression Analysis.....	31
4.2.1 Test of Goodness of Fit of the Model	31
4.2.2 Interpretation of Parameter Estimates.....	32
4.2.3 Model Diagnostic.....	38
CHAPTER FIVE	39
5. Discussion, Conclusion and Recommendations	39
5.1 Discussion and Conclusion	39
5.2 Recommendations.....	41
Reference	43
Appendix.....	47

List of Tables

Table 3.1. List of the names, descriptions and codes of the independent variables	14
Table 4.1. Result of descriptive statistical data analysis.....	27
Table 4.2. Test of Significance of Hosmer-Lemeshow Goodness of Fit Statistic Hosmer-Lemeshow Test	32
Table 4.3. Stata logit output for the standard logistic regression model.....	33

CHAPTER ONE

1. Introduction

1.1 Background of the Study

According to the traditional career development perspective work is seldom just seen as a means by which an individual sustains life (Holland, 1997; Zunker, 1994; Savickas, 1991). Instead, work is viewed as having many dimensions or functions. First, work has an economic function: people work to earn wage, or are involved in activities to be rewarded in such a way that they can sustain themselves and dependents in order to fulfill certain primary needs. Second, there is a social dimension to work. The occupation or work that individuals are involved in determines to a large extent where and how they live, the community and organizations in which they participate and many other social aspects of their lives. Social status has long been associated with individuals' jobs. Third, work has a personal or psychological dimension. It is an essential source of identity, and provides people with the feeling of self-worth and self-esteem as they experience a feeling of mastery and self-fulfillment when they successfully engage in work activities. Research shows that unemployed people often experience feelings of low self-esteem resulting from their not being involved in activities that are valued by other people (Zunker, 1994). The functions of work therefore, are of great importance to both society and the individual.

The discourse on the value of work can only take on real meaning when the issue of unemployment is raised: unemployment has become an immense global problem and it affects a vast majority of the world's population currently. The ILO international standard definition of unemployment is based on the following three criteria which should be satisfied simultaneously: "without work", "currently available for work" and "seeking work".

With a land area of 1.1 million square kilometer and a population of about 73.9 million (50.46 male and 49.54% female) in 2008/09, Ethiopia is the second populous country in Africa next to Nigeria. According to the May 2007 Population and Housing Census of Ethiopia, about 84 percent of the population resided in rural areas agriculture being the major source of livelihood. Although the rate of population growth has been on a declining trend over the last three decades (3 per cent per annum in the 1980s, 2.73 percent up until the early 1990s and 2.6 percent from the mid 1990s up to 2007), Ethiopia's population growth is still considered to be high given its size and demographic profile (CSA, 2008). Ethiopia is a poor agrarian country with per capita income of USD 350 (World Bank, 2011). The proportion of the population living below the poverty line of less than one dollar a day at purchasing power parity is estimated to be 2 per cent for 2000, while 78 per cent of the population lived on less than two dollars a day (again, at purchasing power parity) (MoFED, 2002).

According to the May 2007 Population and Housing Census results (CSA, 2008), Ethiopia's population is predominantly young with about 45% of the population being below 15 years of age. According to the same source, the proportion of working age population (15-64 years) was estimated at about 52 percent. The dependency ratio (number of dependents per 100 working age population) was estimated at 93 by the end of 2007, youth and senior citizens dependency being 87 and 6, respectively (CSA, 2008). High dependency means higher pressure on public services, high level of unemployment, low per capita income, and low level of domestic saving and asset accumulation with serious implication on poverty incidence. This has also serious implication on natural resources degradation with far reaching consequence on sustainable development.

Between 1984 and 2005, the total labor force of the country has more than doubled. It increased from 14.7 million in 1984 to 26.5 million in 1994 and further to 33 million in 2005. Employment creation for such a rapidly increasing labor force (4.4 % per annum), has become increasingly challenging. The labor force participation rate at national level has risen modestly overtime, from 77 percent in 1999 to nearly 82 percent by 2005. The rural participation rate increased from 79 percent in 1999 to 85 percent in 2005 compared to urban participation rate of 63 percent in 1999 and 65 percent in 2005 (LFS, 2005).

1.2 Statement of the Problem

Unemployment is one of the major socio-economic problems. It is one of the greatest economic concerns not only in Ethiopia but also in the world. There are no positive aspects of unemployment and is really bad for the production of goods in the economy. This means that we are wasting our resources because instead of producing goods and services with them we are not doing anything with them. But this is not the only effect of unemployment; people's income will also decrease and it will cause more poverty. Unemployment also reduces the amount of tax paid to the government which means that governments lose money because since less people are working. The ability of governments to provide for people is also seriously compromised. When there is high unemployment, people pay less income tax and also pay less in sales taxes because they purchase fewer goods and services. This leads to less in the way of public services.

If people are unemployed production of goods and provision of services falls off, and simultaneously, the people who are unemployed lack the wherewithal to purchase goods and services. People who still have money, investors, are reluctant to invest any money in the production of goods or the provision of services because when production and consumption are down, there is no opportunity to get a return on the investment.

Effects of unemployment are social too; not just economic. Frequently, crime rates rise as people are unable to meet their needs through work. Divorce rates often rise because people cannot solve their financial problems. The rate of homelessness rises, as do the

rates for mental and physical illness. Homes are foreclosed upon or abandoned, and neighborhoods deteriorate as a result.

In general the effect of unemployment is adverse. Unemployment is not a good thing for anyone in society, and even the people who remain employed will suffer as a result. Ethiopia is one victim of the consequences of unemployment. So studying the determining factors of unemployment and searching for remedies is not negotiable.

1.3 Objective of the Study

General Objective:

- ✓ The general objective of this study is to identify the demographic and socio-economic determinants of unemployment in Ethiopia.

Specific Objectives:

- ✓ To identify the determining factors that are associated with unemployment in Ethiopia.
- ✓ To give relevant recommendations for those concerned.

1.4 Significance of the Study

Currently, (un) employment is a critical concern to almost every country in the world. It is one of the most pressing economic and social problems confronting developing countries whose labor markets have weakened substantially. Many studies have been done on the problem. It is hoped that this study will come up with the major determining factors of unemployment in Ethiopia. The results could be helpful for the formulation of policies and strategies to facilitate the reduction of unemployment in the country. The findings could also be helpful in order to conduct further studies on the issue.

1.5 Limitation of the Study

The limitations of this study are as follows.

- In Ethiopia, there is no national legislation that makes education or schooling compulsory or that forbids children from participating in any production or service activities. Furthermore, except in the public sector, where a minimum age of 18 years was set for entrance into employment, there is no national labor legislation that excludes children from admission to formal or non-formal activities below a certain age (Genene Bizuneh et al, 2001). Participation rate of children 5-14 years is 30% in Sub-Saharan Africa (World Bank, 2008). The current study does not include this age group.
- The study could not include the type of unemployment the individual was facing (cyclical/structural/seasonal/frictional).

CHAPTER TWO

2. Literature Review

Unemployment is not a problem that has emerged recently. It has been a problem for a long time though these days it is getting more and more severe. Hence, different studies have been undertaken throughout the world. In this section some of the findings are discussed.

2.1 Unemployment as a Global Problem

Many studies have similarly argued that marriage and family life were negatively affected by the unemployment experience (Atkinson, Liem and Liem, 1986; Dew, Bromet and Schulberg, 1987; Elder and Caspi, 1988; Jackson and Walsh, 1987; Liem and Liem, 1988; Liem and Rayman, 1982; Moen, 1979; Moen, Kain and Elder, 1983; Nowak and Snyder, 1984; Schlozman and Verba, 1979; Voydanoff and Donnelly, 1988). These works showed conclusively how emotional distress arising from job loss affects both the job loser and other family members.

Some studies have linked unemployment to levels of violence in the family (Straus, Gelles and Steinmetz, 1980). Others have reported an increase in levels of family friction, tension and arguments as a result of unemployment (Grayson, 1985; Hakim, 1982; Komarovsky, 1962). Further, a review documents showed many negative effects of parents' unemployment and associate financial distress on children (McLoyd, 1989).

A study by O'Higgins (2001) reported that although youth unemployment varies from one country to another, a few features were common to most of the nations investigated. First, it was found that youth unemployment was higher than adult unemployment in almost every country for which data was available. In most of these countries youth unemployment was double adult unemployment and, in certain cases, even three times the adult rates. A second common factor of youth unemployment across countries was that it is strongly linked to adult unemployment. It was also found that upsets in the aggregate labor market that have a direct effect on adult unemployment had a more pronounced effect on youth employment. Youth unemployment can therefore, not be separated from the aggregate unemployment situation; the general context will always have an influence on youth unemployment. Lastly, a link was found between (un)employment and economic growth, indicating that output growth is a precondition for employment growth, although the picture for this is clearer in developed countries than in some of the developing countries.

O'Brien (1986) concluded that unemployment experiences are less difficult for unskilled people because they have learnt to expect less of life than people who are better skilled. It was also speculated that youth would be more vulnerable and, therefore, much more affected than adults.

Salvador and Killinger (2008), WB (2009) and Morris (2006) noted that unemployment rate of less educated youth tends to be higher than the unemployment rate of more

educated youth in developing countries because their skills and competencies may not correspond to the demand of the labor market. In other words, the chance of getting employment for more educated youth was higher as compared to lower educated youth since they had the required knowledge and skills. Similarly, Mlatsheni and Rospabe (2002) found that young people with secondary level education (from grade 8 to grade 12) did not have a better chance to get a job than people with no education.

According to Hallerod and Westberg (2006), being one of the demographic variables, sex revealed substantial differences between female and male with respect to employment opportunity. Females were vulnerable both in short term and long term unemployment than males. Strengthening this point, Mlatsheni and Rospabe (2002) found that lack of employment was more severe for females than for males as 63 percent of economically active females were unemployed whereas 53 percent of males remain without jobs in South Africa. They further noted that one of the reasons behind females' unemployment was that girls spend much time in doing domestic work than boys. This leads them to poor academic performance and sometimes withdrawal from education.

Okojie (2003) stated that unemployment in Africa concentrated among youth who have received some education. He further added that youth who had limited education lack the industrial and other skills demanded in the labor market, thereby making them unattractive to employers who prefer skilled and experienced workers. Confirming this idea, Anh et al (2005) found that youth who attained limited education were more prone

to unemployment in the continent. In addition to this, they noted that, training in Africa remains largely unrelated to the labor market needs, which fosters the existence of a degree of mismatch between the demand for and supply of education. Furthermore, Guracello and Rosati (2007) found that the less educated youth face more difficulties in finding employment in urban areas of the country, Ethiopia.

A research conducted by Anh et al (2005) and Rees and Gray (1982) found that family income serves as an important factor in determining the employment experience of Vietnamese youth. A family in which a young person lived was the strongest predictor of his or her future in the job market. On the other side, they added that youth who resided in low income earning family were less employed in the labor market.

Morris (2006) showed that the significant effects of family economic status, paternal occupation, education and parental divorce were notable in affecting the employment status of youth. He further noted that a better income earning household had a number of opportunities, i.e. higher income can enable youth to have greater access to education, information and connections. This could facilitate easy access to employment opportunities available in the market. Strengthening this point, a study conducted by Echebiri (2005) depicted that unemployment had affected youths from a broad spectrum of socioeconomic groups, both the well and less well educated, although it had particularly stricken a substantial fraction of youths from low income backgrounds.

2.2 Unemployment in Ethiopia

Guracello and Rosati (2007) stated that female youth across all ages were more likely to be unemployed and were much more likely to be jobless than male youth in Ethiopia. Another research conducted by Berhanu et al (2005) noted that unemployment rate among young female (20-24) was 38.7 percent while it was only 23.2 percent for young male in the same age category during the same year in Ethiopia. Besides, the CSA (2010a) unemployment report also showed that out of 1,168,591 unemployed persons 41.2 percent were female youth. Furthermore, Bizuneh et al (2001) confirmed that females were more marginalized than males due to different socioeconomic factors. Hence, the problem of unemployment was more prevalent among females than males.

Nzinga H. Broussard and Tsegay Gebrekidan Tekleselassie studied about youth unemployment in Ethiopia based on the data from the 1999/2000 and 2004/2005 labor force surveys, and the 2009 and 2011 urban employment and unemployment surveys. Their report provides a comprehensive description of the main characteristics of the youth labor market in Ethiopia. They found that while unemployment in urban areas remains widespread, it declined markedly since 1999 for the economy as a whole and for youth. However, while the economy has demonstrated impressive reductions in unemployment, women have not benefited as much as men. They had significantly higher unemployment rates than their male counterparts and were often confined to the informal sector. In Ethiopia, there has been significant increase in educational attainment;

however, there has not been as much job creation to provide employment opportunities to the newly educated job seekers.

A number of studies have looked at different aspects of the urban labor market in Ethiopia (Krishnan et al., 1998; Serneels, 2001; Bizuneh et al., 2001). Findings from these studies indicate the very high level of unemployment in urban Ethiopia. Based on the 1994 census, Bizuneh et al (2001) stated that the level of urban unemployment was 30 percent for men and 40 per cent for women in Addis Ababa, and about 15 percent for both men and women in other urban centers in 1994. They also reported that the overwhelming majority of the unemployed were made up of first time job seekers, emphasizing the problem that the urban youth finds itself in. Focusing on the 20 – 29 age group in Addis Ababa, they reported that the general unemployment rate for men stood at 50 per cent while it was 60 per cent for women.

In his unemployment duration study that focused on young men, Serneels (2001) also stated the magnitude of the unemployment problem that the youth faced. He stated that in 1994 urban Ethiopia had one of the highest unemployment rates in the world standing at 34 per cent of the male workforce and 50 per cent of men under 30 years of age.

CHAPTER THREE

3. Data and Methodology

3.1 Data Source

The data was taken from the Ethiopia Demographic and Health Survey (EDHS) conducted by Central Statistics Agency (CSA) in 2011. The 2011 Ethiopia Demographic and Health Survey (2011 EDHS) was conducted under the aegis of the Ministry of Health (MOH) and was implemented by the Central Statistical Agency from September 2010 through June 2011 with a nationally representative sample of nearly 18,500 households. All women age 15-49 and all men age 15-59 in these households were eligible for individual interview. The 2011 EDHS sample was selected using a stratified, two-stage cluster design, and enumeration areas (EAs) were the sampling units for the first stage. The 2011 EDHS sample included 624 EAs, 187 in urban areas and 437 in rural areas.

3.2 Description of Variables

Both the dependent and the independent variables were selected based on available similar studies. The independent variables were thought to be determining factors of the response variable, that is, unemployment.

3.2.1 The Dependent Variable

The 2011 EDHS asked respondents whether they were currently employed at the time of the survey (that is, had worked in the past seven days).

The dependent variable is Unemployment Status of an Individual in Ethiopia. The response variable is dichotomous. If the i^{th} individual is unemployed, the response variable (Y_i) takes the value 1 otherwise it takes the value 0.

$$Y_i = \begin{cases} 1, & \text{if the } i^{\text{th}} \text{ individual is unemployed} \\ 0, & \text{if the } i^{\text{th}} \text{ individual is employed} \end{cases}$$

3.2.2 Independent Variables

The independent variables that are used in this paper are given in the table below. As determinants of unemployment, this study utilized only socio-economic and demographic factors.

Table 3.1. List of the names, descriptions and codes of the independent variables

Variables	Descriptions	Codes
SEX	Sex of the respondent	1=Female 2=Male
AGE	Age of the respondent	1=15-24 2=25-39 3=Above 39
REG	Region	1=Tigray 2=Affar

		3=Amhara 4=Oromia 5=Somali 6=Benishangul-Gumuz 7=SNNP 8=Gambela 9=Harari 10=Addis Ababa 11=Dire Dawa
RES	Place of residence	1=Urban 2=Rural
EDNL	Educational level	0=No education 1=Primary 2=Secondary 3=Higher
SEXHHH	Sex of head of household	1=Male 2=Female
EMM	Exposure to mass media	0=Not at all 1=Less than once a week 2=At least once a week
ECOST	Economic status	1=Poor 2=Medium 3=Rich
MARST	Marital status	0=Never in union 1=No longer living with partner 2= Married

HHS	Household size	1=Small (1 to 4 members) 2=Medium (5 members) 3=Medium large (6 or 7 members) 4=Large (8 or 9 members) 5=Very large (>9 members)
NCH	Number of living children	0=No child 1=Small (1 or 2 child/ren) 2=Medium (3 children) 3=Medium large (4 or 5 children) 4=Large (6 or 7 children) 5=Very large (>7 children)

For the variable EMM, in Table 3.1, the mass media includes reading newspaper, listening to the radio and watching TV.

3.3 Methodology

3.3.1 Logistic Regression

Logistic regression is a type of regression analysis used for predicting the outcome of a categorical (a variable that can take on a limited number of categories) dependent variable based on one or more predictor variables. The probabilities describing the possible outcome of a single trial are modeled, as a function of explanatory variables, using a logistic function.

Logistic regression measures the relationship between a categorical dependent variable and a set of predictor variables. Logistic regression can be binomial or multinomial.

Binomial or binary logistic regression refers to the instance in which the observed outcome can have only two possible types. Logistic regression is used when the dependent variable is dichotomous or polytomous.

Logistic regression is used to predict the odds of being a case based on the predictor(s). The odds are defined as the probability of a case (unemployment) divided by the probability of a non case (employment) and is given by:

$$\text{Odds} = \frac{\pi}{1-\pi} \quad (3.1)$$

where π is the probability of success and $1-\pi$ is the probability of failure. Logistic regression is used in a wide range of applications leading to binary dependent data analysis.

3.3.1.1 Assumptions of the Binary Logistic Regression

- Logistic regression does not assume a linear relationship between the dependent and independent variables.
- The dependent variable must be binary (2 categories) and does not need to be normally distributed, but it typically assumes a distribution from an exponential family (e.g. binomial, Poisson, multinomial, normal); binary logistic regression assumes binomial distribution of the response.

- The categories (groups) must be mutually exclusive and exhaustive; a case can only be in one group and every case must be a member of one of the groups.
- Larger samples are needed than for linear regression because maximum likelihood coefficients are large sample estimates. A minimum of 50 cases per predictor is recommended.

3.3.1.2 The Model

In the terminology of logistic regression analysis the odds of a success is defined to be the ratio of the probability of a success to the probability of a failure.

Let Y be an $n \times 1$ vector of response variable with $Y_i = 1$ if the person is unemployed and $Y_i = 0$ if the person is employed, X is an $n \times (k+1)$ design matrix of explanatory variables and β is a $(k+1) \times 1$ vector of parameters to be estimated.

Let $\pi(x)$ denotes the conditional probability that the outcome is unemployment (probability of success).

$$\begin{aligned}\pi(x) &= P(Y=1/X) = 1 - P(Y=0/X) \\ &= \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)} = \frac{\exp(X' \beta)}{1 + \exp(X' \beta)}\end{aligned}\quad (3.2)$$

We obtain the odds of success by using the above equation.

$$\text{odds}(Y=1) = \frac{\pi(x)}{1 - \pi(x)} = \exp(X' \beta) \quad (3.3)$$

A transformation of $\pi(x)$ that is central to our study of logistic regression is the logit transformation. This transformation is defined, in terms of $\pi(x)$, as:

$$\text{logit}(\pi) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \log(\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)) = \log(\exp(X'\beta)) \quad (3.4)$$

The importance of this transformation is that it has many of the desirable properties of a linear regression model. The logit function is linear in its parameters, may be continuous, and may range from $-\infty$ to $+\infty$, depending on the range of x .

The transformed variable, denoted by $\text{logit}(\pi)$ is the log-odds and is related to the explanatory variables as:

$$\text{logit}(\pi) = \eta(x) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k = X'\beta \quad (3.5)$$

where $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)'$ are the model parameters and $X' = (X_0, X_1, \dots, X_k)$ with $X_0 = 1$, are explanatory variables.

In logistic regression we may express the value of the outcome variable given x as $y = \pi(x) + \varepsilon$. Here the quantity ε may assume one of two possible values. If $y = 1$ then $\varepsilon = 1 - \pi(x)$ with probability $\pi(x)$, and if $y = 0$ then $\varepsilon = -\pi(x)$ with probability $1 - \pi(x)$. Thus, ε has a distribution with mean zero and variance equal to $\pi(x)[1 - \pi(x)]$. That is, the conditional distribution of the outcome variable follows a binomial distribution with probability given by the conditional mean, $\pi(x)$ (Hosmer and Lemeshow, 2000).

3.3.1.3 Estimation of Model Parameters

The most commonly used method of estimating the parameters of a logistic regression model is the method of Maximum Likelihood (ML) instead of Ordinary Least Squares (OLS) method. Mainly for this reason the ML method based on Newton-Raphson iteratively reweighted least square algorithm becomes more popular with researchers (Ryan, 1997).

In a very general sense the method of maximum likelihood yields values for the unknown parameters which maximize the probability of obtaining the observed set of data. In order to apply this method we must first construct a function, called the likelihood function. This function expresses the probability of the observed data as a function of the unknown parameters. The maximum likelihood estimators of these parameters are chosen to be those values that maximize this function. Thus, the resulting estimators are those which agree most closely with the observed data (Hosmer and Lemeshow, 2000).

3.3.1.4 Goodness of Fit of the Model

Goodness of fit in linear regression models is generally measured using the R^2 . Since this has no direct analog in logistic regression, various methods including the following can be used instead (Hosmer and Lemeshow, 2000).

I. Deviance and Likelihood Ratio Tests

In linear regression analysis, one is concerned with partitioning variance via the sum of squares calculations – variance in the criterion is essentially divided into variance accounted for by the predictors and residual variance. In logistic regression analysis, deviance is used in lieu of sum of squares calculations. The statistic, D , in equation (3.6) is what is called the deviance; according to authors [McCullagh and Nelder (1989)]. It plays a central role in some approaches to assessing goodness-of-fit. Deviance is analogous to the sum of squares calculations in linear regression and is a measure of the lack of fit to the data in a logistic regression model. Deviance is calculated by comparing a given model with the saturated model – a model with a theoretically perfect fit. This computation is called the likelihood-ratio test.

$$D = -2 \ln \left[\frac{\text{likelihood of the fitted model}}{\text{likelihood of the saturated model}} \right] \quad (3.6)$$

In the above equation D represents the deviance and $\ln$ represents the natural logarithm. The results of the likelihood ratio (the ratio of the fitted model to the saturated model) will produce a negative value, so the product is multiplied by negative two times its

natural logarithm to produce a value with an approximate chi-squared distribution (Hosmer and Lemeshow, 2000). Smaller values indicate better fit as the fitted model deviates less from the saturated model. When assessed upon a chi-square distribution, non significant chi-square values indicate very little unexplained variance and thus, good model fit. Conversely, a significant chi-square value indicates that a significant amount of the variance is unexplained.

Two measures of deviance are particularly important in logistic regression: null deviance and model deviance. The null deviance represents the difference between a model with only the intercept and no predictors and the saturated model. And, the model deviance represents the difference between a model with at least one predictor and the saturated model. In this respect, the null model provides a baseline upon which to compare predictor models. Given that deviance is a measure of the difference between a given model and the saturated model, smaller values indicate better fit. Therefore, to assess the contribution of a predictor or set of predictors, one can subtract the model deviance from the null deviance and assess the difference on a chi-square distribution with one degree of freedom. If the model deviance is significantly smaller than the null deviance then one can conclude that the predictor or set of predictors significantly improved model fit. This is analogous to the F-test used in linear regression analysis to assess the significance of prediction (Cohen, Cohen, West and Aiken, 2002).

II. Hosmer–Lemeshow Test

The Hosmer–Lemeshow test uses a test statistic that asymptotically follows a χ^2 distribution to assess whether or not the observed event rates match expected event rates in subgroups of the model population. The test hypothesis is given by:

H_0 : The model fits the data. Vs.

H_1 : The model does not fit the model.

The test statistic is constructed by grouping the data set into groups (often $g=10$). The groups are formed by ordering the existing data by the level of their predicted probabilities. So the data are first ordered from least likely to have the event to most likely for the event. Then g roughly equal sized groups are formed. The observed and expected numbers of events are computed for each group. The test statistic is:

$$\hat{C} = \sum_{K=1}^g \frac{(O_k - E_k)^2}{v_k} \quad (3.7)$$

where, O_k and E_k are the observed and expected number of events in the k^{th} group, and v_k is a variance correction factor for the k^{th} group. If the observed number of events differs from what is expected by the model, the statistic $\hat{C}$ will be large and there will be evidence against the null hypothesis. This statistic has an approximate chi-squared distribution with $(g - 2)$ degrees of freedom. The advantage of a summary goodness-of-fit

statistic like $\hat{C}$ is that it provides a single, easily interpretable value that can be used to assess fit (Hosmer and Lemeshow, 2000).

3.3.1.5 Test of Significance of an Individual Predictor

I. Likelihood Ratio Test

The likelihood-ratio test discussed above to assess model fit is also the recommended procedure to assess the contribution of individual "predictors" to a given model (Menard, 2002). In the case of a single predictor model, one simply compares the deviance of the predictor model with that of the null model on a chi-square distribution with a single degree of freedom. If the predictor model has a significantly smaller deviance (compare chi-square using the difference in degrees of freedom of the two models), then one can conclude that there is a significant association between the "predictor" and the outcome.

II. Wald Statistic

Alternatively, when assessing the contribution of individual predictors in a given model, one may examine the significance of the Wald statistic. The Wald statistic, analogous to the t-test in linear regression, is used to assess the significance of coefficients. The Wald statistic is the ratio of the regression coefficient to the square of the standard error of the coefficient and is asymptotically distributed as a chi-square distribution (Menard, 2002). The Wald statistic used to test the hypothesis $H_0: \beta_j = 0$ against $H_1: \beta_j \neq 0$ is:

$$W_j = \frac{\hat{\beta}_j^2}{(SE(\hat{\beta}_j))^2} \quad (3.8)$$

Although several statistical packages report the Wald statistic to assess the contribution of individual predictors, the Wald statistic is not without limitations. When the regression coefficient is large, the standard error of the regression coefficient also tends to be large increasing the probability of Type-II error. The Wald statistic also tends to be biased when data are sparse (Cohen, Cohen, West and Aiken, 2002).

3.3.1.6 Model Diagnostic

Before concluding that the model "fits", it is crucial that other measures be examined to see if fit is supported over the entire set of covariate patterns. This is accomplished through a series of specialized measures falling under the general heading of regression diagnostics.

Model diagnostic procedures involve both graphical methods and formal statistical tests. These procedures allow us to explore whether the assumptions of the regression model are valid and decide whether we can trust subsequent inference results.

I. DFBETAs: assess the effect of an individual observation on the estimated parameter of the fitted model. A DFBETAS diagnostic is computed for each observation for each parameter estimate. It is the standardized difference in the parameter estimate due to deleting the corresponding observation. The DFBETAs are useful in detecting observations that causes instability in the selected coefficients.

II. Leverage (hat matrix): an observation with an extreme value on the predictor variable is called a point with high leverage. Leverage is a measure of how far an observation deviates from the mean of that variable. These leverage points can have an effect on the estimate of regression coefficients.

III. Cook's distance (D): measures of how much the residual of all cases would change if a particular case were excluded from the calculation of the regression coefficients. A large Cook's distance indicates that excluding a case from computation of the regression statistics changes the coefficients substantially (Cook and Weisberg, 1982).

CHAPTER FOUR

4. Statistical Data Analyses

4.1 Descriptive Data Analysis

The household and individual response rates for the 2011 EDHS are reported in the EDHS preliminary report. A total of 17,817 households were selected for inclusion in the 2011 EDHS, and of these, 17,018 were included in the study. Of the 17,018 households, 16,702 were successfully interviewed, yielding a response rate of 98 percent. In the interviewed households, a total of 17,385 women were identified to be eligible for the individual interview, and 95 percent of them were successfully interviewed. For men, 15,908 were identified as eligible for interview, and 89 percent of them were successfully interviewed. The table below shows the descriptive statistics for the selected variables. SPSS was applied for this analysis.

Table 4.1. Result of descriptive statistical data analysis

		UNEMPLOYED			
		No		Yes	
		Count	Percentage	Count	Percentage
SEX	Female	5933	36.0	10540	64.0
	Male	7533	55.5	6041	44.5
AGE	15-24	2666	22.4	9252	77.6
	25-39	6835	54.7	5666	45.3
	Above 39	3965	70.5	1663	29.5
REG	Tigray	1245	40.8	1805	59.2
	Afar	758	33.8	1486	66.2
	Amhara	1753	44.3	2205	55.7
	Oromiya	2040	49.2	2104	50.8

	Somali	608	38.1	988	61.9
	Benishangul-Gumuz	1226	52.2	1123	47.8
	SNNP	1732	46.8	1971	53.2
	Gambela	923	45.9	1087	54.1
	Harari	936	46.2	1088	53.8
	Addis Ababa	1380	45.9	1625	54.1
	Dire Dawa	865	44.0	1099	56.0
RES	Urban	4244	45.5	5078	54.5
	Rural	9222	44.5	11503	55.5
EDNL	No education	5967	47.7	6541	52.3
	Primary	5259	42.7	7043	57.3
	Secondary	1121	38.1	1824	61.9
	Higher	1119	48.8	1173	51.2
HHS	Small	1376	45.8	1626	54.2
	Medium	1675	47.4	1856	52.6
	Medium large	4256	48.7	4489	51.3
	Large	3603	44.5	4502	55.5
	Very large	2556	38.4	4108	61.6
NCH	No child	3461	26.8	9450	73.2
	Small	7323	51.0	7026	49.0
	Medium	795	95.1	41	4.9
	Medium large	1107	96.3	42	3.7
	Large	507	97.3	14	2.7
	Very large	273	97.2	8	2.8
SEXHHH	Male	10969	46.5	12606	53.5
	Female	2497	38.6	3975	61.4
EMM	Not at all	4553	40.5	6696	59.5
	Less than once a week	4435	46.3	5138	53.7
	At least once a week	4478	48.5	4747	51.5
ECOST	Poor	4742	43.5	6168	56.5
	Medium	2021	46.5	2325	53.5
	Rich	6703	45.3	8088	54.7

MARST	Never in union	1602	15.9	8443	84.1
	No longer living with partner	1017	53.6	879	46.4
	Married	10847	59.9	7259	40.1

The proportion of unemployment status of individuals varied based on sex in Ethiopia.

Higher number of unemployed individual was observed for the women (64.0%). Men had 44.5% of unemployment rate.

The highest proportion of unemployment was observed in the age group 15-24 (77.6%).

The individuals in the age group 25-39 had the second highest proportion of unemployment, that is, 45.3%. Individuals older than 39 years had 29.5% of unemployment rate.

Besides, the above table shows that the proportion of unemployment differed from one region to another. The greatest number of unemployed individuals was recorded in Afar region (66.2%). The regions Somali and Tigray followed Afar with 61.9% and 59.2%, respectively. The smallest number of unemployed individuals was recorded in Benishangul-Gumuz (47.8%). The regions Oromiya, SNNP and Harari followed Benishangul-Gumuz with 50.8%, 53.2% and 53.8%, respectively. Regions Amhara and Dire Dawa had unemployment rate of 55.7% and 56.1%, respectively. Addis Ababa and Gambela had equal rate of unemployment (54.1%).

The rate of unemployment in urban areas was a little bit different from the rate in rural areas. The proportion of unemployment was 54.5% in urban areas while it was 55.5% in rural areas.

Similarly, the rate of unemployment also differed based on the education level of the individual. Some 52.3% of the individuals who had no education were unemployed. The proportion of unemployment for the individuals who had primary and secondary school were 57.3% and 61.9%, respectively. The lowest percentage of unemployment was observed in individuals who completed higher education (51.2%).

The rate of unemployment in small household size was 54.2%. About 52.6% of the individuals from a medium family size were unemployed. 51.3%, 55.5% and 61.6% of the individuals from medium large, large and very large household sizes, respectively, were found to be unemployed.

Higher proportion of unemployment was observed in households where women were head (61.4%) while relatively smaller percentage (53.5%) of unemployment was observed in households where men were head.

The rate of unemployment decreased with increasing frequency of exposure to the mass media. About 59.5% of the respondents who did not have exposure at all were

unemployed. Some 53.7% and 51.5% of the respondents who had exposure to mass media less than once a week and at least once a week were unemployed, respectively.

About 56.5% of the poor individuals were unemployed. The rich and the medium income individuals had 54.7% and 53.5% of unemployment rates, respectively.

The highest percentage (73.2%) of unemployment was for individuals with no children while the lowest percentage (2.7%) was for individuals with large number of children. Individuals with small, medium, medium large and very large number of children had 49.0%, 4.9%, 3.7% and 2.8% of unemployment rate, respectively.

The highest percentage (84.1%) of unemployment was observed among individuals who were never in union. Some 40.1% of the individuals who were married were unemployed at the time of the survey. Out of the individuals who were no longer living with their partner, 46.4% were unemployed.

4.2 Logistic Regression Analysis

4.2.1 Test of Goodness of Fit of the Model

Logistic regression uses maximum likelihood, which is an iterative procedure. The first iteration (called iteration 0) is the log likelihood of the "null" or "empty" model; that is, a model with no predictors. At the next iteration, the predictor(s) are included in the model. At each iteration, the log likelihood increases because the goal is to maximize the

log likelihood. When the difference between successive iterations is very small, the model is said to have "converged".

The deviance (-2LL) is a measure of the difference between a given model and the saturated model, smaller values indicate better fit. The model deviance is significantly smaller than the null deviance; hence the set of predictors significantly improved model fit.

Table 4.2. Test of Significance of Hosmer-Lemeshow Goodness of Fit Statistic Hosmer-Lemeshow Test

Chi-square	Df	Sig.
14.80	8	0.0632

The p-value (0.0632) which is greater than 0.05 in the above table gives an evidence not to reject H_0 . Hence the model is a good fit to the data.

4.2.2 Interpretation of Parameter Estimates

In this section binary logistic regression is applied to assess the relation between unemployment status of a person, which is a dichotomous response variable, with the explanatory variables. STATA version 11.0 is used to perform the binary logistic regression analysis. Table 4.3 displays the maximum likelihood estimates of parameters where the significance of each parameter is tested using the Wald test.

Table 4.3. Stata logit output for the standard logistic regression model

UNEMPLOY	Coef.	Std. Err.	z	P> z	Odds Ratio	[95% Conf. Interval]	
						LB	UB
1.SEX	.699	.066	21.25	0.000*	2.011	1.886	2.145
AGE			206.87	0.000*			
1	.617	.090	12.75	0.000*	1.853	1.685	2.037
2	.154	.047	3.80	0.000*	1.166	1.077	1.262
REG			367.73	0.000*			
1	-.006	.070	-0.08	0.935	.994	.866	1.142
2	.469	.123	6.08	0.000*	1.599	1.374	1.860
3	-.222	.055	-3.26	0.001*	.801	.701	.915
4	-.401	.045	-5.96	0.000*	.669	.587	.764
5	.380	.121	4.58	0.000*	1.462	1.243	1.720
6	-.529	.044	-7.14	0.000*	.589	.510	.681
7	-.310	.051	-4.47	0.000*	.733	.640	.840
8	-.114	.068	-1.50	0.135	.892	.768	1.036
9	-.092	.068	-1.24	0.215	.912	.788	1.055
10	-.404	.047	-5.79	0.000*	.668	.582	.765
1.RES	-.309	.034	-6.63	0.000*	.734	.670	.804
EDNL			137.40	0.000*			
0	.649	.123	10.07	0.000*	1.914	1.687	2.172
1	.332	.081	5.69	0.000*	1.393	1.243	1.562
2	.490	.109	7.34	0.000*	1.632	1.432	1.861
HHS			123.08	0.000*			
1	-.445	.036	-7.91	0.000*	.641	.574	.716
2	-.412	.034	-7.92	0.000*	.663	.598	.734
3	-.412	.028	-9.81	0.000*	.662	.610	.719
4	-.196	.035	-4.55	0.000*	.822	.756	.894
NCH			3.00	0.223			
0	3.320	9.074	10.11	0.000*	27.647	14.530	52.607
1	2.813	5.450	8.60	0.000*	16.655	8.770	31.630
2	.990	.939	2.83	0.005*	2.690	1.357	5.333

3	.484	.567	1.38	0.167	1.622	.817	3.219
4	.232	.497	0.59	0.556	1.261	.583	2.728
1.SEXHHH	.464	.060	12.42	0.000*	1.591	1.479	1.712
EMM			1.12	0.143			
0	.158	.046	4.08	0.000*	1.172	1.086	1.264
1	-.072	.034	-1.96	0.050	.931	.866	1.000
ECOST			14.78	0.006*			
1	.145	.047	3.60	0.000*	1.156	1.068	1.251
2	.035	.050	0.74	0.461	1.036	.943	1.138
MARST			1440.01	0.000*			
0	1.829	.302	37.70	0.000*	6.229	5.664	6.851
1	.663	.101	12.76	0.000*	1.940	1.752	2.148
_cons	-4.277	.338	-12.67	0.000*		-4.939	-3.616

Note: The symbol * indicates that the estimate is significant at 0.05. Reference categories are: 'men' for SEX 'above39' for AGE, 'Dire Dawa' for REG 'rural' for RES, 'higher' for EDNL, 'women' for SEXHHH, 'rich' for ECOST, 'very large' for HHS and 'married' for MARST.

In Table 4.3, the column labeled “coef.” is the estimated values of the parameters for the logistic regression equation. They are in log-odds units. Because these coefficients are in log-odds units, they are often difficult to interpret, so they are often converted into odds ratio.

The values in “Std.Err.” column, in Table 4.3, are the standard errors associated with the coefficients. The standard error is used for testing whether the parameter is significantly different from 0; by dividing the parameter estimate by the standard error we obtain a z-value (see the column with z-values and p-values). The standard errors can also be used

to form a confidence interval for the parameter, as shown in the last two columns of this table.

Column-6 of the table gives the odds ratios for each variable. The last categories of the explanatory variables are used as reference. The odds ratio is the ratio of the odds of an event occurring in one group to the odds of it occurring in another group

The odds of unemployment of women was 2.011 (OR=2.011) times higher than the odds of unemployment of men controlling other variables in the model.

Individuals in the age group 15-24 were 85.3% more likely to be unemployed (OR=1.853) compared to individuals in the age group greater than 39 years; while individuals in the age group 25-39 were 16.6% (OR=1.166) more likely to be unemployed compared to individuals in the age group greater than 39 years controlling the other variables in the model.

Individuals who lived in Afar were 59.9% (OR=1.599) more likely to be unemployed compared to Dire Dawa controlling for the other variables in the model. The odds of unemployment decreased by a factor of 0.801 (OR=0.801) and 0.669 (OR=0.669) when the person was from Amhara and Oromia, respectively, compared to a person from Dire Dawa. The odds of unemployment in Benishangul-Gumuz, Gambela, Addis Ababa and SNNP were 0.589 (OR=0.589), 0.892 (OR=0.892), 0.668 (OR=0.668) and 0.733

(OR=0.733) times lower than the odds of unemployment in Dire Dawa, respectively. The odds of unemployment in Harari and Tigray were not significantly different from the odds of unemployment in Dire Dawa controlling for the other variables in the model. Individuals who lived in Somali were 46.2% more likely to be unemployed (OR=1.462) compared to individuals who lived in Dire Dawa.

The odds of unemployment of individuals who lived in urban areas was 0.734 (OR=0.734) times lower than the odds of unemployment of individuals who lived in rural areas controlling the other variables in the model.

Individuals who had no education were 91.4% (OR=1.914) more likely to be unemployed compared to individuals who had higher education controlling for other variables in the model. The odds of unemployment of individuals who had primary education was 1.393 (OR=1.393) times higher than the odds of unemployment of individuals who had higher education; while the odds of unemployment of individuals who had secondary education was 1.632 (OR=1.632) times higher than the odds of unemployment of individuals who had higher education controlling for the other variables in the model.

Individuals who lived in households where men were head were 59.1% (OR=1.591) more likely to be unemployed compared to individuals who lived in households where women were head of the household controlling for the other variables in the model.

Individuals who did not have exposure to mass media at all were 17.2% (OR=1.172) more likely to be unemployed compared to individuals who had the exposure at least once a week. The odds of unemployment of those who had exposure to mass media less than once a week was not significantly different from those who had the exposure at least once a week controlling for the other variables in the model.

Individuals in small household size were 0.641 (OR=0.641) times less likely to be unemployed compared to individuals in a very large household size; while individuals in medium, medium large and large household sizes were 0.663 (OR=0.663), 0.662 (OR=0.662) and 0.822 (OR=0.822) times less likely to be unemployed, respectively, compared to individuals in a very large household size controlling for the other variables in the model.

The odds of unemployment of individuals who lived in poor households was higher by a factor of 1.156 (OR=1.156) compared to the odds of unemployment who lived in rich households. There was no significant difference between the odds of unemployment of individuals who lived in medium and rich-income households.

The odds of unemployment for individuals who were never in union was 6.229 (OR=6.229) times higher than those who were married. Individuals who were no longer living with partner were 94% (OR=1.940) more likely to be unemployed compared to individuals who were married.

4.2.3 Model Diagnostic

After fitting a regression model it is important to determine whether all the necessary model assumptions are valid before performing inference. If there are any violations, subsequent inferential procedures may be invalid resulting in faulty conclusions. Therefore, it is crucial to perform appropriate model diagnostics. The results of the model diagnostic are shown in the Appendix.

All the values of the model diagnostic measures are less than unity. A DFBETAs less than unity indicates that no one value has an effect on the estimate of a regression coefficient of a particular predictor variable. Cook's distance less than unity implies that no one observation has an effect on a group of regression coefficients. A value of the leverage statistic less than unity confirms that no observation is far apart from the others in terms of the levels of the independent variables. Therefore, the model diagnostic measures prove model adequacy.

CHAPTER FIVE

5. Discussion, Conclusion and Recommendations

5.1 Discussion and Conclusion

This study had the objective to identify the most important socio-economic and demographic determinants of unemployment based on the 2011 Ethiopia Demographic and Health Survey. According to the descriptive data analysis, Afar had the highest percentage of unemployment. Somali and Tigray followed. The lowest percentage was observed in Benishangul-Gumuz. The variation might be the result of way of living or cultural effect. In some cultures women are not allowed to engage themselves in a formal work sector; this increases the number of unemployed. Unlike the findings in Serneels (2001) and Bizuneh et al. (2001), the higher rate of unemployment was observed in rural areas of the country than urban. A substantial difference in unemployment between women and men with respect to employment opportunity was observed. This is consistent with the findings of Bizuneh et al. (2001) which confirmed that, the problem of unemployment was more prevalent among females than males. Strengthening this point, Mlatsheni and Rospabe (2002) found that lack of employment is more severe for women than for men. Consistent with Okojie (2003), the lowest number of unemployed was seen for those who had higher education; while the highest was for individuals with secondary education. The reason behind this is, the fact that higher education provides the skills needed to perform complex jobs, making people more productive, thus sustaining economic growth. People with the most human capital are said to be the most productive,

and thus secure the best jobs and the highest salaries. Thus, education plays an important role in determining the employment status of an individual. This study showed that individuals in the age group 15-24 were found to be more affected than the other age groups. Similar findings were obtained by O'Brien (1986) and O'Higgins (2001) confirming that youth (age 15-24) face more difficulties in finding employment. Specifically the current study found that women in the age group 15-24, across all ages, were more likely to be unemployed. Guracello and Rosati (2007) also confirmed that female in the age group 15-24 face more difficulties in finding employment in Ethiopia. Individuals who resided in poor economy were more vulnerable than those who lived in medium or rich economy. Morris (2006) noted that better income earning households had a number of job opportunities and individuals who lived in households with low income were less employed in the labor market.

In conclusion, according to the binary logistic regression result, the predictors age, sex, region, place of residence, educational level, economic status, marital status, sex of head of household and household size ($p < 0.05$ in all cases) were found to have a significance effect on unemployment status of an individual in Ethiopia.

5.2 Recommendations

The story that comes out of this study is that unemployment in Ethiopia is very high. This is an obvious story that probably does not deserve these many pages of telling. What is probably more important is to come up with some suggestions that will hopefully help alleviate the problem of unemployment in Ethiopia. In the following part, some attempt will be made towards achieving this goal.

There are different things that can be done in the long run and short run. In the short term creating jobs that require manual work by humans is possible. In this way even the less qualified individuals will get the opportunity of being employed. Empower women and increase their participation; as women are more vulnerable to unemployment. Efforts made by the government and other organizations to do so should further be enhanced.

In the long run, increasing and improving the level of education will create better chances not to live in poverty. Though unemployment is the result of aggregate socio-economic factors, education is the best way of reducing it. Beside, creation of new jobs can be achieved through education.

Finally, the task of fighting unemployment in Ethiopia is unlikely to be an easy one. It is a task that government alone cannot carry out. A concerted effort from all the major actors is absolutely vital. The state should be at the forefront of the fight by creating an enabling environment and by carrying out the much-needed task of coordination. The

private sector, NGOs, the donor community, and other important actors should be there to complement any such effort.

Reference

- Anh, D., Duong, L. and Van, N. (2005). Youth Employment in Vietnam: Characteristic, Determinants and Policy Responses. Employment Strategy Papers. ILO, Geneva.
- Atkinson, T., Liem, R., and Liem, J. H. (1986). The social costs of unemployment: Implications for social support. *Journal of Health and Social Behavior*, 27, 317-331.
- Berhanu Denu, Abraham Tekeste and Hannah van der Deijl (2005). Characteristics and determinants of youth unemployment, underemployment and inadequate employment in Ethiopia: Employment Strategy Paper.
- Bizuneh, G., Adino, T., Gesano, G., Guarneri, A. and Heins, F. (2001). Work status and unemployment in urban Ethiopia, mimeo, Central Statistical Authority (Addis Ababa, Ethiopia) and Institute for Population Research – National Research Council (Roma, Italy).
- Cohen, J.; Cohen, P.; West, Steven G.; Aiken, Leona S. (2002). *Applied Multiple Regression Correlation Analysis for the Behavioral Sciences* (3rd ed.). Routledge.
- Cook, R. and Weisberg, S. (1982). *Residuals and Influence in Regression*. Chapman and Hall, New York.
- CSA (2008). Summary and Statistical Report of the 2007 Population and Housing Census. Population Size by Age and Sex. Addis Ababa, Ethiopia.
- CSA (2010a). Analytical Report on the 2009 Urban Employment and Unemployment Survey. Addis Ababa, Ethiopia.
- Dew, M. A., Bromet, E. J., and Schulberg, H. C. (1987). A comparative analysis of two community stressors' long-term mental health effects. *American Journal of Community Psychology*, 15, 167-184.
- Echebiri, R.N. (2005). Characteristics and Determinants of Urban Youth Unemployment in Umuahia, Nigeria: Implications for Rural Development and Alternative Labor Market Variables. Okpara University of Agriculture, Umuahia, Abia State.

- Elder, G. H., Jr., and Caspi, A. (1988). Recent research on the psychological effects of unemployment. *Journal of Social Issues*, 44(4), 25-45.
- Ethiopia Demographic and Health Survey (2011). Central Statistical Agency Addis Ababa, Ethiopia ICF International Calverton, Maryland, USA.
- Grayson, P.J. (1985). The closure of a factory and its impact on health. *International Journal of Health Services*, 15, 69-93.
- Guarcello, L. and Rosati, F. (2007). Child Labor and Youth Employment: Ethiopia Country Study. SP Discussion Paper No.0704. The World Bank.
- Hakim, C. (1982). The social consequences of high unemployment. *Journal of Social Policy*, 2, 433-467.
- Hallerod, B. and Westberg, A. (2006). Youth Problem: What is the Problem?: A Longitudinal Study of Incomes and Economic Hardship among Swedish Youth. Sage Publications, Ltd., *Sociological Review*, Vol. 49, No. 1, PP (83-102).
- Holland, J. L. (1997). Making vocational choices: A theory of vocational personalities and work environments (3rd ed.). Odessa FL: Psychological Assessment Resources.
- Hosmer, W.D. and S. Lemeshow (2000). *Applied Logistic Regression*. 2nd Ed., John Wiley and Sons, New York.
- Jackson, P. R., and Walsh, S. (1987). Unemployment and the family. In D. Fryer & P. Ullah (Eds.), *Unemployed people: Social and psychological perspectives*. Philadelphia: Open University Press.
- Komarovsky, M. (1962). *Blue collar marriage*. New York: Random House.
- Krishnan, P., T. Gebreselassie and S. Dercon (1998). The urban labor market during structural adjustment: Ethiopia 1990 – 1997. Centre for the Study of African Economics.
- LFS (2005). Central Statistical Agency, Addis Ababa.
- Liem, R., and Liem, J. H. (1988). Psychological effects of unemployment on workers and their families. *Journal of Social Issues*, 44(4), 87-105.
- Liem, R., and Rayman, R. (1982). Health and social costs of unemployment: Research and policy considerations. *American Psychologist*, 37, 1116-1123.

- McCullagh, P., and Nelder, J. A. (1989). *Generalized Linear Models*, Second Edition. Chapman and Hall, London.
- McLoyd, V. C. (1989). Socialization and development in a changing economy. *American Psychologist*, 44, 293-302.
- Menard, Scott W. (2002). *Applied Logistic Regression* (2nd ed.). SAGE.
- Mltasheni, C. and Rospabe, S. (2002). Why is Youth Unemployment so High and Unequally Spread in South Africa? School of Economics University of Cape Town, Working Paper 02/65.
- Moen, P. (1979). Family impacts of the 1975 recession: Duration of unemployment. *Journal of Marriage and the Family*, 41, 561-572.
- Moen, P., Kain, E., and Elder, G. H. Jr. (1983). Economic conditions and family life: Contemporary and historical perspectives. In R. Nelson and F. Skidmore (Eds.), *American families and the economy*. Washington, DC: National Academic Press.
- MoFED (2002). *Sustainable Development and Poverty Reduction Program* Federal Democratic Republic of Ethiopia, Addis Ababa, Ethiopia.
- Morris, E. (2006). *Globalization and Its Effects on Youth Employment Trends in Asia*. ILO Sub Regional Office for East Asia. Geneva.
- Nowak T. C., and Snyder, K. A. (1984). Job loss, marital happiness and household tension: Do women fare better than men? Paper Presented at the Society for the Study of Social Problems Annual Meeting. San Antonio, TX.
- Nzinga H. B. and Tsegay G.T. (2012). *Youth Unemployment: Ethiopia Country Study*.
- O'Brien, G.E. (1986). *Psychology of work and unemployment*. New York: Wiley.
- O'Higgins, N. (2001). Youth unemployment and employment policy: a global perspective. MPRA Paper 23698, University Library of Munich, Germany.
- Okojie, C. (2003). *Employment Creation for Youth in Africa: The Gender Dimension*. Economics and Statistics University of Benin City, Nigeria.
- Rees, A. and Gray, W. (1982). *The Youth Labor Market Problem: Its Nature, Causes and Consequences*. University of Chicago Press, Chicago.

- Ryan, T. (1997). *Modern Regression Methods*. Wiley, Inc., New York.
- Salvador, R., and Killinger, N. (2008). *An Analysis of Youth Unemployment in the Euro Area*. Occasional Paper Series, No. 89, Frankfurt.
- Savickas, M. L. (1991). The meaning of work and love: Career issues and interventions. *The Career Development Quarterly*, 39, 315-324.
- Schlozman, K. L., and Verba, S. (1979). *Injury to insult: Unemployment, class and political response*. Cambridge, MA: Harvard University Press.
- Serneels, P.(2001). *Unemployment duration in urban Ethiopia: Centre for the Study of African Economics*.
- Straus, M. A. Gelles, R. J., and Steinmetz, S. K. (1980). *Behind Closed Doors: Violence in the American Family*. Garden City, NY: Doubleday, Anchor Press.
- Voydanoff, P., and Donnelly, B. W. (1988). Economic Distress, Family Coping, and Quality of Family Life. In P. Voydanoff and L. C. Majka (Eds.), *Families and Economic Distress: Coping Strategies and Social Policy*. Newbury Park, CA: Sage.
- World Bank (2011). *World Development Indicators*.
- World Bank (2008). *World Development Report 2008: Agriculture for Development*. Washington, DC.
- World Bank (2009). *Africa Development Indicator: Youth and Employment in Africa*. Washington, D.C., U.S.A.
- Zunker, V.G. (1994). *Career Counseling: Applied Concepts of Life Planning*. Fourth edition, Belmont, CA: Brooks/Cole Publishing.

Appendix

Summary of descriptive statistics for outliers and influential observations diagnostics

Descriptive Statistics

	N	Minimum	Maximum
DFBETA for constant	30396	-.19571	.16669
DFBETA for SEX(1)	30396	-.00135	.00062
DFBETA for AGE(1)	30396	-.00209	.00155
DFBETA for AGE(2)	30396	-.00143	.00147
DFBETA for REG(1)	30396	-.00284	.00272
DFBETA for REG(2)	30396	-.00293	.00284
DFBETA for REG(3)	30396	-.00298	.00273
DFBETA for REG(4)	30396	-.00283	.00264
DFBETA for REG(5)	30396	-.00351	.00342
DFBETA for REG(6)	30396	-.00278	.00302
DFBETA for REG(7)	30396	-.00276	.00273
DFBETA for REG(8)	30396	-.00292	.00323
DFBETA for REG(9)	30396	-.00263	.00224
DFBETA for REG(10)	30396	-.00272	.00241
DFBETA for RES(1)	30396	-.00188	.00186
DFBETA for EDNL(1)	30396	-.00338	.00277
DFBETA for EDNL(2)	30396	-.00313	.00258
DFBETA for EDNL(3)	30396	-.00265	.00217
DFBETA for NHHM(1)	30396	-.00271	.00236
DFBETA for NHHM(2)	30396	-.00251	.00211
DFBETA for NHHM(3)	30396	-.00238	.00210
DFBETA for NHHM(4)	30396	-.00312	.00330
DFBETA for SEXHHH(1)	30396	-.00129	.00102
DFBETA for LIT(1)	30396	-.16723	.19555
DFBETA for LIT(2)	30396	-.16719	.19628
DFBETA for LIT(3)	30396	-.16730	.19648
DFBETA for LIT(4)	30396	-.16690	.19629
DFBETA for FREQRNP(1)	30396	-.00286	.00255
DFBETA for FREQRNP(2)	30396	-.00244	.00218
DFBETA for FREQLR(1)	30396	-.00118	.00118

DFBETA for FREQLR(2)	30396	-.00083	.00106
DFBETA for FREQWTV(1)	30396	-.00186	.00170
DFBETA for FREQWTV(2)	30396	-.00139	.00163
DFBETA for ECOST(1)	30396	-.00111	.00131
DFBETA for ECOST(2)	30396	-.00164	.00166
DFBETA for NLCH(1)	30396	-.01129	.00367
DFBETA for NLCH(2)	30396	-.01081	.00369
DFBETA for NLCH(3)	30396	-.01051	.00387
DFBETA for NLCH(4)	30396	-.01034	.00389
DFBETA for NLCH(5)	30396	-.01014	.00685
DFBETA for MARST(1)	30396	-.00181	.00155
DFBETA for MARST(2)	30396	-.00205	.00281
Valid N (listwise)	30396		

Descriptive Statistics

	N	Minimum	Maximum
Analog of Cook's influence statistics	30396	.00001	.16525
Valid N (listwise)	30396		

Descriptive Statistics

	N	Minimum	Maximum
Leverage value	30396	.00032	.05903
Valid N (listwise)	30396		

DECLARATION

I the undersigned, declare that the thesis is my original work, has not been presented for degree in any other University and that all sources of material used for the thesis have been duly acknowledged.

Declare by

Name _____

Signature _____

Date _____

Place of submission: Department of statistics, college of natural science, Addis Ababa University.

This thesis has been submitted for examination with my approval as a university advisor.

Name _____

Signature _____

Date _____