COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCE

DEPARMENT OF STATISTICS

ASENIOR RESEARCH PAPER ON THE DETERMINANTS OF URBAN YOUTH UNEMPLYMENT INETHIOPIA

PREPARED BY

1. AHMED AWOL……………………………. …….NCSR/ 029/09
2. TAMRU BERTA…………………………………...NCSR/ 530/09

ASENIOR RESEARCH PAPER SUMITTED TO THE DEPARTMENT OF STATISTICS IN PARTAIL FULFLUMENT OF THE REQUIRMENT FOR THE DEGREE OF BATCHELOR OF SCIENCE IN STATISTICS

Advisor: Abu Tuna (MS.c)

JUNE, 2019

WOLKITE, ETHIOPIA

ACKNOWNLEDGEMENT

First of all, we would like to thank GOD for enabling us to undertake this research*unlimited support throughout our life. Next to this, our deepest gratitude goes to our advisor Mr. Abu Tuna(M.SC) his relevant suggestion and continuous provision of constructive idea which brought this practical research to compilation and to build our performance. Finally we would also like to thanks all staff members of Statistics department who directly or indirectly participated for the accomplishment of this paper and all our department students.*

ABSTRACT

This study is an attempt to investigate the factors that determine youth unemployment status in Ethiopia. The main objective of this study was to identify determinants unemployment status in Ethiopia. The data were analyzed by statistical software packages such as SPSS. The study uses both descriptive and inferential statistics method of data analysis, which means from descriptive frequency table and from inferential statistics chi-square test of associations and binary logistic regression. In chi square test of association variables such as region,education level, sex,age marital statues have association with employment status. The variables education level, region and drug addictionhad statistically significant effect in logistic regression on unemployment status. This study recommended that the concerning body should have to decreases this effects of unemployment we must decreases the unemployment number by creating job opportunity to all citizen.

List of Abbreviations

CSA **Central Statistical Agency**

ICLS **International Conference of Labor Statisticians**

ILO **International Labor Organization**

LF **Labor force**

LFS **Labor Force Survey**

LL **Likelihood**

LR **Likelihood Ratio**

ML **Maximum Likelihood**

OR **Odds Ratio**

S.E **Standard Error**

SPSS **Statistical Package for Social Science**

List of tables

Table of content

CHAPTER ONE

1. 1. Background of the Study

Unemployment is one of the main problems in the world economy today. Many countries at different levels of development are trying to cope with this problem. Unemployed not working and not contributing for income generation of the economy, but also they claim benefits from the government and it is an additional cost for the economy as a whole [7]. The government defines those who want to work as people who have actively looked for work within the past four weeks and determines the number of people currently unemployed through a monthly survey called the Current Population Survey

People can be unemployed for many reason such as quit their position and are looking for a new one, TheirCompany reduced the work force, and they are seeking a new position. This can be due to a local condition, when the company closes a plan or division, or a national condition, when the economy slows and many companies reduce their work force, They have recently returned to the work force -perhaps from pregnancy or attending school and haven't yet located a position. ,the need for their skill set has gone down, and there are limited positions available, which may lead to unemployment until they train for a new position and technology has reduced the need for their type of position.

International Labor Organization (1992) defines that unemployment is the situation of being out of work or need a job and continuously searching for it in the last four week or unemployed (age 18 or above) but available to join work in the next two weeks. A person to be categorized as unemployed, two conditions must be fulfilled: That the person is without a job and able to workand The person wants to have a job and searching to work at the current market wage ratePeople who voluntarily do not want to work, full time students, retired people and children are noincluded in unemployed category. In short, unemployment means the state when people who are willing and able to do a job but fail to get the desired job.

Youth unemployment is a problemthat affects most countries. Theability of youthto engage in productive activities has both social and economic consequences for an economy.

Youth unemployment is often higher than the unemployment rate for adults highlighting the concerns that many countries face in facilitating the transition from school to work. In developing countries, youth face not only the challenge of obtaining productive employment, butalso obtaining safe and acceptable works[20].

Government organizations, NGOs and civic association in different countries adopt and use various age ranges for the concept "youth" from the stand point of the purpose which they stand for and the activities they undertake. For Example, the United Nation (UN) defines the youth as persons between 15 -24 years; WHO, 10-24. In Ethiopia, according to the national youth policy, youth includepart of the society who is between 15-29 years [18].

Unemployment in Ethiopia is more of a problem of urban youth than that of rural. According to Ethiopian labor force survey report, the unemployment rate of urban youth at country level were 22.9 while for rural youth remained at 3.1% only[17].

Based on the above facts, this study was tried to analyze the determinants of urban youth unemployment in Ethiopia.

1.2Statement of the problem

Today unemployment is one of the most challenges and very serious issues facing in Africa and also in the World.Like many other developing countries unemployment has been one of the major problems in Ethiopia. The excessive rate of unemploymentnegatively impacts on economy which causes unstable economic conditions. This is troublesome because when workers are unemployed, there is an under-utilization of resources. So the total production of a country is less than its potential level of output because resources are not fully utilized in these countries [19].

A high level of underemployment is one of the critical socio-economic problems facing Ethiopia. While the labor force grows with an increasing proportion of youth, employment growth is inadequate to absorb labor market entrants. As aresult, youth are especially affected by unemployment. Moreover, young people are more likely to be employed in jobs of low quality, underemployed, working long hours for low wages, engaged in dangerous work or receive only short term and informal employment arrangements [8]

Unemployment is of a special concern for Ethiopians and has a wider implication for the youth in addition to leading their life asexpected to help parents and extended families[21]. According to a survey in 55 urban areas, unemployment was estimated at 41.3% and the incidence of youth unemployment was 45.5% and 35.7% for females and males respectively [3]. Even though few studies are conducted on youth unemployment in Ethiopia the results of these studies contradict each other whichneed further study based on the specific socio-economic situation of the study area. Hence, this study will be conducted on the determinants of urban youth unemployment in Ethiopia.

1.3 Objective of the study

1.3.1 General objective of the study

- The general objective of this study is to analyze the determinants of urban youth unemployment in case ofurban Ethiopia.

1.3.2 Specific objective of the study

- ✓ To assess the overalldeterminants of urban youth unemployment using descriptive statistics.
- ✓ To investigate the associationbetween determinants of urban youth unemployment andage, gender, marital status, region and educational level

1.4 Significance of the study

This study was beneficial to address the problem of urban youthunemployment and identifying its effect in Ethiopia. After conceder their effects the all concerned body take action by may assisting the people in finding the jobs. By inviting different private and governmental investors to the area then the problem will be solving. This study also helps to provide basic information for the concerned with unemployment and helps them to find the main problem that related with study will be giving a base line data for future study and to address the major effect and related consequence of unemployment problemof the study area and give recommendations to concerned bodies to decreases the problem of unemployment in the cases of Ethiopia. It helps for the researcher to improve the employment and solving related problems simply.

1.5 Scopes of the study

The scope of the study was recovered on someurbanEthiopia, to identify the determinants of unemployment.

CHAPTER TWO

2. Literature Review

The ILO estimates that the number of unemployed youth is on the rise again since 2011, after declining from the peak it reached at the height of the global financial crisis. It is expected to reach 74.2 mi

llion young people by 2014 based on [16]. The global youth unemployment rate has also been rising since 2011. It is currently estimated at 12.6% and is projected to increase to 12.8% by 2018.

The study by[12], on the basis of youth in Umuahia city in Nigeria, finds that unemployment is influenced by age, marital status, education, current income and employment preference (paid or self-employment)[4]also indicate that age, gender, marital status, region and educational level are the major determinants of unemployment in Jordan.

Studies from Ethiopia indicate that the potential causes of unemployment in urban Ethiopia include increasing number of youth labor force, the rising internal migration, literacy rate, poor to modest macroeconomic performance, low level of job creation and low level of aggregate demand in the economy[13,24] ; Youth unemployment is the outcome of different socio-economic and demographic factors at macro and micro level. The micro level factors are directly associated to individuals' demographic and socioeconomic attributes while the macro level factors are related to the national issues [23].

This study emphasizes on assessing individuals' demographic and socioeconomic attributes that influence youth employment[6], the multivariate analysis showed that sex, , education, job preferences and access to business advisory services significantly determine youth unemployment in DebreBirhan town.

However, maritalstatus werefound insignificantly related to youth unemployment.According to[22], examined the association between socio-demographic variables and unemployment in Addis Ababa, the econometric analysis has confirmed that sex and age are statistically significant and have negative relationship, signifying the inherent problem of unemployment among women and the youth. Regarding migration status, in spite of the type of job, a migrant is more likely to be employed than a non-migrant. This result can be an indication of the obvious fact that there is unmet demand for domestic and casual labor in the city, a pull factor for the rural poor and marginalized youth, particularly women. Thus, given the existing push and pull factors from rural areas and the unmet labor demand in urban centers; the migrants' supply of labor would be mutually beneficial to both the urban as well as the rural communities.[16], conducted the binary logistic regression to assess the determinants of youth unemployment at Ambo, Ethiopia. Their result showed that among the demographic variables, age of the respondents and migration status were significantly related to youth unemployment whereas marital status of the respondents was not significant. From the human capital variables included in the model, education and health status of the respondents were significantly related to youth unemployment, whereas participation in employment related trainings was not statistically significant. Among the economic determinants, household income, access to credit and saving services and work experience were significant. Access to job information and psycho-social factors were the two social capital variables that were significantly related to youth unemployment. As youths are more vulnerable to unemployment, efforts should be made by the government to provide credit and training so as to facilitate their entry into business and entrepreneurship. Migrants are the victims of unemployment in town. Therefore, the pushing factors of migrants should be identified to arrest the continuous drift of youth towards urban areas as this may worsen the unemployment situation in urban areas.

CHAPTER THREE

3. METHODOLOGY

3.1Description of the Study Area

The study was conducted on urban youth unemployment in case of Ethiopia. Ethiopia is one of the developing countries from Africa and found near to east Africa. It is bordered by Eritrea to the north, Djibouti and Somali to the east, Sudan to the west, and Kenya to the south. Ethiopia has a high central plateau that varies from 1,290 to 3,000m above sea level , with the highest mountain reaching 4,533m the area of Ethiopia is 1.13 million sq km(437,794 sq miles and 102.3 million population distributed .It has nine regional state and two administration state.

3.2 Methods of data collection

Basically we have two main categories of data source mainly primary source of data and secondary source of data. Depending on the source, Primary method, consist of obtaining data or information by any one of the following ways: direct personal interview, in direct oral interview, information from correspondents, mailed questionnaire and so on. On the other hand, secondary data would be obtained from document; books and the governmental sector in the study area and collected from different publication, like annual reports of the central statistics agency of Ethiopia. But our study was dependedon secondary data

3.3 Study population

The target populations of this study was the total number of youth population that is registered in urban Ethiopia $.i.e$Youthpopulationincluding age from (18- 30)

3.4 Study variables

3.4.1 Dependant variable

Dependant variable: The variable we wish to explain (also called the endogenous variables).The response (Dependant) variable of this study is unemployment status of urban youth in Ethiopia

$$Y = \begin{cases} 1, \text{urban youth is unemployed} \\ 0, \text{urban youth is employed} \end{cases}$$

3.4.2 Independent variables

Independent variables: The variable used to explain the dependant variable (also called exogenous variable). Or a variable that is influenced the dependent variable. In this studyindependent variables are;

Table-1: Summary of independent variables that may influence on the dependent variable

Variable	Description	values
Sex	Sex of the respondent	0 = female and 1= male
Age	age of the respondent	Continuous variable
Marital status	Marital status of the respondent	0=married
		1=single
		2=widowed
Education level	Educational level of the respondent	0= primary education
		1= secondary education
		2=preparatory
		3=illiterate
		4= Higher education
Drug addiction	Drug addiction of the respondent	1=no
		2=yes

3.5 Method of Data Analysis

To analyze the collected data, the study wasused both

- Descriptive and
- Inferential statistics.

3.5.1 Descriptive Statistics;

Descriptive statistical analysis summary measure such as number and proportion would be used to describe the sample information. It is utilized numerical and graphical methods to look for patterns in the data set to summarize and to present that information in convenient form. It describes the data collected through charts, frequency table, Statistical graphs and so on. In this study, frequency tables are used.

3.5.2 Inferential statistics;

It deals with making inference or conclusion about population based on data obtains from a limited number of observations from population. In this study we employed both chi-square test of independency and binary logistic regression model to analyze the data.

3.5.2.1 Chi-square test of Independence

The chi square test of independence is used to test the association between two variables. It is used to analysis categorical data. The chi- square independent test is use to whether there is association between dependent and independent variable (Bluman, 3rd Edition)

Here, $\chi^2 = \sum\sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_\alpha((R-1)(C-1))$ 3.1

χ^2cal=chi square calculated

O_{ij}= observed frequency.

E_{ij}= expected frequency.

The test statistic is distributed approximately as chi-square with (r-1) (c-1) degree of freedom. Note that: c is number of the columns on the data sand r is number of rows on the data.

Expected frequency $= \frac{\sum_{i=1}^{n} R \sum_{j=1}^{n} C}{N}$

Where R is row total and C is column total

3.5.2.1.1 Hypothesis test

Chi - square test is applicable for testing the association between two variables. The null and the alternative hypothesis will be states as

Ho: There is no association between two variables. i.e. there is no association between unemployment and predicted variable.

H1: not Ho.

Decision Rule: - Reject the null hypothesis and do not reject the alternative hypothesis, if p-value is less than level of significance.

Assumptions of chi- square test of independence

- The populations must be normally distributed for the variable under the study.
- The observations must be independent each other.
- The sample must be randomly selected from the population.
- The sample size is large.
- Each data cell should contain at least five observations
- The expected frequency for each category must be 5 or greater than 5

3.5.2.2Logistic regression analysis

Logistic regression is applied when the dependent variable is qualitative in nature or categorical. Qualitative variable are either binary (dichotomous) variable or multiply categories. Binary logistic regression is the form of regression which is used when the dependent variables dichotomous and the independent variable are of any type. A logistic regression is a technique for making prediction when the dependent variable is dichotomous and the independent may be categorical and mix of continuous categorical (Christensen, 1997).

3.5.2.2.1 Binary logistic regression model

It is used to when the response or dependant variable is categorical. The independent variable may be qualitative or categorical. Due to the categorical nature of dependent variable the technique of binary logistic regression is used for analysis. The logistic regression method can be used not only to identify the risk factors, but to predict the probability of success. The logistic regression applies maximum likelihood estimation after transforming the dependant in to logitvariable (the natural log of the odd of the dependant variable occurring or not). .

The dependent variable in this case is dummy variable, which take the value of (1) for u$nemployed$ and (0) foremploye.

(i.e. . .) Yi= $\begin{cases} 0, \text{employed} \\ 1, \ \text{u}nemployed \end{cases}$

Assumption of binary Logistic Regression Model

- Normally distributed errors terms are not assumed.
- Logistic regression does not assume a linear relationship between the dependent and the Independentvariables.
- The dependent (outcome) variables are dichotomous and the independent variables are either categorical or continuous variables
- Linearity in the logit regression equation should have a linear relationship with the logit form of be the dependent variable.
- Absence of multi co linearity.
- Logistic regression requires large sample to guarantee higher level of accuracy.
- The dependent variables need not be homoscedasticity for each level of independent variables; that is there is no homogeneity of variance assumption

A model with one or more predictors is fit using an iterative reweighted least squares algorithm to obtain maximum likelihood estimates of the parameters. Binary logistic regression has also been used to classify observations in to one of two categories and it may give fewer classification errors than discriminates analysis for some cases. The model for the binary logistic regression is given as:

$\{\frac{p}{1-p}\} = exp^{\beta 0+\beta 1x1+\beta 2x2+}\ ... + \beta kxk$……………………………3.2

Ln $\{\frac{p}{1-p}\} = \beta 1 + \beta 1x1 + \beta 2x2 + \cdots + \beta kxk$………………………3.3

Where,

- ✓ P= the probability woman non user on of maternal health service.
- ✓ 1-p= the probably woman user on of maternal health service.
- ✓ βo= is constant term.
- ✓ β_is are coefficient of the predictor variable
- ✓ X_i, i=1, 2, 3 is independent variables.

So by using SPSS we are going to obtain values for each coefficients because the model have been Programmed in to available logistic regression packages like SPSS .

Estimating the Parameters in Logistic Regression Model

The maximum likelihood parameters estimation method used to estimate parameters of logistic regression model.

3.6 StatisticalModel

The model with binary logistic response variable in regression problem takes only two possible values, 1 = unemployed, 0=employed.

The relationship between the predictor and response variables is not a linear function in logistic regression; instead of, the logistic regression function, which is the logit transformation of π, is used. Consider a collection of explanatory variables denoted by the vector $x'=(x_1, x2\ldots x_p)$.let the conditional probability that the outcome is success be denoted by p (y=1/x) =π

$$\pi=\frac{\exp(\beta0+\beta1x1+\cdots+\beta pxp)}{1+\exp(\beta0+\beta1x1+\cdots+\beta pxp)}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.4)}$$

Then the logit or log- odds of having y=1is a model as a linear function of the explanatory variables

as: $Ln(\frac{\pi}{1-\pi})=\beta0+\beta1x1+\ldots.+\beta pxp;\ 0\leq \pi \leq 1\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.5)}$

Where: -β0is the constant of the equation andβ1,...,are the coefficients of the predictorvariables.

3.6.1 Odds Ratio

Logistic regressions work with odds so it is necessary to define both odds and odds ratio. The odds are simply the ratio of the probabilities for the two possible outcomes. If p is the probability that the event will occur, then 1 –p is the probability that the event will not occur: odds=$\frac{p}{1-p}$in 2 × 2 tables, within row 1 the odds of success are odds1=$\frac{p_1}{1-p_1}$and within row 2 the odds of success equal odds2=$\frac{p_2}{1-p_2}$.The ratio of the odds from the two rows$pi =$

$\frac{odds1}{odds2}$ is called odds ratio. Whereasthe relative risk is a ratio of two probabilities, the odds ratio $p_i$is a ratio of two odds. Interpretation of odds ratio are this$e^{\beta}j$ is the factor by which the odds changes when the jth independent variable increase by one unit. Note: if β_j is positive then the odds increase and if β_j is negative, the odds decrease.

3.6.2Parameter Estimation for Logistic Regression

To estimate the parameters of logistic regression model, the two estimation methods mostly used are maximum likelihood and non-iterative weighted least squares method. When the assumption of normality of the predictors does not hold, the non- iterative weighted least squares method is less efficient. In contrast, the maximum likelihood estimation method is appropriate for estimating the logistic model parameters due to this less restrictive nature of the underlying assumptions. Thus in this study the maximum likelihood estimation technique will be applied to estimate parameters of the model. Consider the logistic regression model$p(xi) = \frac{e^{xi\beta}}{1+e^{xi\beta}}$. Since observed values of Y say, Y_i's (i=1, 2… n) are independently distributed as Bernoulli random variable, the probability mass function of Yi is:

$(yi) = \ (1-\pi i)^{yi}$……………………………………………………….. (3.7)

Yi = 0 or 1 and i=1, 2…, n

Since the Yi are assumed to be independent, the likelihood function is given by:

$$L(\beta / y) = \prod_{i=1}^{n} P(y_i \mid X_i^{'}) = \prod_{i=1}^{n} \left[\frac{e^{X_i^{'}\beta}}{1+e^{X_i^{'}\beta}} \right]^{y_i} \left[\frac{1}{1+e^{X_i^{'}\beta}} \right]^{(1-y_i)}$$ 3.8

The principle of maximum likelihood states that we use as our estimate of the value which maximizes the likelihood function. However, it is easiermathematically to work with the log likelihood function. the log likelihood is:$L(\beta_0,\ \beta_1,\ \dots\ ,\ \beta_p)$=$\sum_{i=1}^{n} yi(\beta_0+\beta_1 x_1+\ldots+\beta_p x_p)$-$\sum_{i=1}^{n} ln\{1 + \exp\ (\beta_0+ \beta_1 x_1+ \cdots + \beta_p x_p)\}$…………..3.9

The objective of stating likelihood function is to get an estimator$\acute{\beta}$= ($\acute{\beta}0$, $\acute{\beta}_1 \ldots \acute{\beta}k$) of$\beta$ which maximizes the likelihood function. In fact all the parameters$\acute{\beta}_0$, $\acute{\beta}_1 \ldots \acute{\beta}_k$and estimates of$(yi|X)$ for each subject could be facilitated by the widely available statistical software .

3.7 Assessment of Model Adequacy

After the model is fitted the next important step is checking the model adequacy. There are several steps involved in assessing the appropriateness, adequacy and usefulness of the model. First, the overall goodness of fit of the model is tested. Second, the importance of each of the explanatory variables is assessed by carrying out statistical tests of the significance of the coefficients.

3.7.1 Goodness of Fit of the model

The goodness of fit of a model measures how well the model describes the response variable. Assessing goodness of fit involves investigating how close values predicted by the model are to the observed values. The appropriateness of the fitted logistic regression model needs tobe examined before it is accepted for use as in the case of all regression models. In practice, several different measures exist for determining the significance or goodness of fit of a logistic regression model. These are Pearson, Hosmer-Lemeshow, Deviance goodness of fit, likelihood ratio test and the classification table. In theoretical sense, all measures are equivalent.

To be more precise, as the number of observation goes to infinity, all measures converge to the same estimate of the model significances. The test can detect major departures from a logistic response function (Alan, 1990).

3.7.2 Test of overall model fit

3.7.2.1Likelihood-Ratio Test

The likelihood ratio test statistic (G^2) is the test statistic commonly used for assessing the overall fit of the logistic regression model[1] argues that the likelihood ratio test is better, particularly if the sample size is small or the parameters are large. The likelihood-ratio test uses the ratio of the maximized value of the likelihood function for the full model (L1) over the maximized value of the likelihood function for the simpler model (L0).

The likelihood-ratio test statistic is given by:

$G^2 = -2\log\left(\frac{L0}{L1}\right) = -2[\log(L_0) - \log(L_1)] = -2(L_0 - L_1)$……………… (3.10)

Where: - L_1 is the likelihood of the full model and

L_0 is the likelihood of the null model.

The likelihood ratio test statistic has an approximate distribution with k degrees of freedom. (Where, k is the number of predictors in thefull model). If significant, it suggests that, taken together, the predictors contribute significantly to the prediction of the outcome. It tests the null hypothesis that all population in logistic regression coefficients is zero except the constant one. And a small p-value, for example, $p < 0.05$ leads to rejection of the null hypotheses that all of the predictor effects are zero. Thus when likelihood test is significant, at least one of the predictors is significantly related to the response variable.

3.7.2.2 The Deviance

The 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 .The null deviance represents the difference between a model with only the intercept (which means "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 and assess the difference on a x^2_{s-p} chi-square distribution with degree of freedom equal to the difference in the number of parameters estimated.

$$D_{null} = -2\ln\left(\frac{likelihoodofnullmodel}{likeliofthesaturatedmodel}\right), \quad D_{fitted} = -2\ln\left(\frac{likelihoodoffittedofmodel}{likelihoodofthesaturatedmodel}\right)$$

$$D = D_{null} - D_{fitted} = -2\ln\left(\frac{likelihoodofnullmodel}{likelihoodoffittedmodel}\right) \ldots\ldots 3.11$$

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.

3.7.2.3Hosmer and Lemshow

The final measure of model fit is the Hosmer-Lemeshow goodness of fit statistics, which measure the correspondence between the actual and the predicted value of the dependent variables. The Homers -Lemeshow test is commonly used test for assessing the goodness of fit model and allows for any number of explanatory variables, which may be continuous or categorical. In this case better model fit is indicated by smaller difference in observed and predicted statistic. The Hosmer–Lemeshow test uses a test statistic that asymptotically follows a x^2distribution to assess whether or not the observed event rates match expected event rates in subgroups of the model population.

The test statistic is: $G^2HL=\sum_1^k(\frac{(o_k-EK)^2}{EK(1-\frac{EK}{NK})})$……..………..……………………………(3.12)

Where, Ok and Ek are the observed and expected number of events in the k^{th} group, and 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 G^2_{HL} 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 G^2_{HL} is that it provides a single easily interpretable value that can be used to assess fit [14].

Ho: The model is good to fit the data.

H1: The model is not good to fit the data.

If the Hosmer-Lemeshow goodness-of-fit test statistic is greater than 0.05,we will not reject the null hypothesis that there is no difference between observed and model predicted values; implying that the model estimates are adequate to fit at acceptable level.

3.7.3Test of individual predictor

3.7.3.1The Wald Statistic

The Wald statistic is an alternative test, which is commonly used to test the significance of individual logistic regression coefficients for each independent variable (that is to test the null hypothesis in logistic regression model that a particular logit coefficient is zero). If the Wald test is not significant, then these explanatory variables can be omitted from the model.

Wald $\chi 2$statistic is used to test the significance of individual coefficients in the model and is calculated as:W= $\left(\frac{\hat{\beta}_j}{se(\hat{\beta}_j)}\right)^2$ ....3.13)where: $-\hat{\beta}i$is the estimated parameter for $\beta i$and SE ($\hat{\beta}i$) is standard error of$\hat{\beta}\iota$

Each Wald statistic is compared with a $\chi 2$distribution with 1 degree of freedom. Wald statistic is easy to calculate but their reliability is questionable, particularly for small samples [2]. For data that produce large estimates of the coefficient, the standard error is often inflated, resulting in a lower value of the Wald statistic, and therefore the explanatory variable may be incorrectly assumed to be unimportant in the model. Likelihood ratio tests are generally considered to be superior.

Pearson's chi-squared test (χ^2) is a statistical test applied to sets of categorical data to evaluate how likely it is that any observed difference between the sets arose by chance. It is suitable for unpaired data from large samples. It is the most widely used of many chi-squared tests (e.g., Yates, likelihood ratio, portmanteau test in time series, etc.) – statistical procedures whose results are evaluated by reference to the chi-squared distribution. Its properties were first investigated by Karl Pearson in 1900. In contexts where it is important to improve a distinction between the test statistic and its distribution, names similar to *Pearson χ-squared* test or statistic are used.

3.8 Strategies in model selection

Model selection for logistic regression faces the same issues as for ordinaryregression. The selection process becomes harder as the number of explanatory variables increases, because of the rapid increase in possible effects andinteractions. There are two competing goals: The model should be complexenough to fit the data well. On the other hand, it should be simple tointerpret, smoothing rather than overfitting the dataMany model selection procedures exist, no one of which is always best.

3.8.1 Stepwise Procedures

In exploratory studies, an algorithmic method for searching among modelscan be informative if we use results cautiously. Goodman 1971a proposedmethods analogous to forward selection and backward elimination in ordinary regression.

The method of forward selection proceeds as follows.

1. Begin with no terms in the model.

2. Find the term that, when added to the model, achieves the largest value of the log likelihood. Enter this term into the model.

3. Continue adding terms until a target value for the log-likelihood is achieved or until a preset limit on the maximum number of terms in the model is reached.

Note that these terms can be limited to those keeping the model hierarchical. This method is comparatively fast, but it does not guarantee that the best model is found except for the first step when it finds the best single term. You might use it when you have a large number of observations and terms so that other, more time consuming, methods are not feasible.

Backward elimination begins with a complex model and sequentiallyremoves terms. At each stage, it selects the term for which its removal has the least damaging effect on the model e.g., largest P-value. The processstops when any further deletion leads to a significantly poorer fit. With eitherapproach, for qualitative predictors with more than two categories, theprocess should consider the entire variable at any stage rather than justindividual dummy variables. Add or drop the entire variable rather thanjust one of its dummies. Otherwise, the result depends on the coding. Thesame remark applies to interactions containing that variable.

Many statisticians prefer backward elimination over forward selection, feeling it safer to delete terms from an overly complex model than to addterms to an overly simple one. Forward selection can stop prematurelybecause a particular test in the sequence has low power. Neither strategynecessarily yields a meaningful model. Use variable selection procedures withcaution!

. 3.8.2 AIC, Model Selection

Other criteria besides significance tests can help select a good model interms of estimating quantities of interest. The best known is the Akaike information criterion AIC. It judges a model by how close its fitted values tend to be to the true values, in terms of a certain expected value. Even though a simple model is farther from the true model than is a more complex model, it may be preferred because it tends to provide better estimates of certain characteristics of the true model, such as cell probabilities. Thus, the optimal model is the one that tends to have fit closest to reality. Given a sample, Akaike showed that this criterion selects the model that minimizes

AIC=2(maximized log likelihood - number of parameters in model). This penalizes a model for having many parameters. With models for categorical Y, this ordering is equivalent to one based on an adjustment of the deviance, [G^2-2(df)], by twice its residual df. For cogent arguments supporting this criterion, see Burnham and Anderson 1998.

CHAPTER FOUR

4. RESULT AND DISCUSSIONS

In this chapter, we are going to analyze the determinants of urban youth unemployment inEthiopia. The data used in this study for the analysis were obtained from CNS EmploymentUN employment data in Ethiopian collected in 2018.

Survey (CNS) with reference to a total of13, 628 youth in the age group 18-30 years.

The dependent variable is a dichotomous random variable —unemployment (coded as 1) and employment (coded as 0). Descriptive and binary logistic regression methods are used to measure the determinants of urban youth unemployment inEthiopia.

The descriptive part provides percentages of total urban youth unemployment towards the predictor variables. The chi- Square analyses used to test the association between the outcome variableand predictor variable.The binary logistic regression analysis used toassess the determinate of urban youth unemployment and predict there effect of the predictor variable . The data are analyzed using the Statistical Package for Social Sciences (SPSS) version 13.

4.1Descriptive statistics

Table 4.1: Summary of descriptive statistics

Variable	Category	Employment status		Total
		Employment	unemployment	
Region	Tigray	705(72.9%)	262(27.1%)	967(100%)
	Afar	342(70.1%)	146(29.9%)	488(100%)
	Amahar	1901(71.7%)	749(28.3%)	2650(100%)
	Ormoiya	2549(70.4%)	1070(29.6%)	3619(100%)
	Somali	402(73.0%)	149(27.0%)	551(100%)
	Benishangul-Gumuz	357(77.4%)	104(22.6%)	461(100%)
	SNNPR	1126(70.4%)	473(29.6%)	1599(100%)
	Gambela	376(71.6%)	149(28.4%)	525(100%)
	Harari	322(73.5%)	116(26.5%)	438(100%)
	AddisAbeba	1372(72.3%)	526(27.7%)	1898(100%)
	Deridewa	320(74.1%)	112(25.9%)	432(100%)
Sex	Male	5177(69.4%)	2287(30.6%)	7464(100%)

	Female	4595(74.5%)	1569(25.5%)	6164(100%)
Education	Primary	2761(68.6%)	1264(31.4%)	4025(100%)
	Secondary	2522(68.3%)	1172(31.7%)	3694(100%)
	Preparatory	301(68.9%)	136(31.1%)	437(100%)
	Illiterate	10(14.7%)	58(85.3%)	68(100%)
	Higher education	4178(77.3%)	1226(22.7%)	5404(100%)
Marital status	Married	3928(65.8%)	2045(34.2%)	5973(100%)
	Single	5410(77.6%)	1559(22.4%)	6969(100%)
	Widowed	434(63.3%)	252(36.7%)	686(100%)
Drug addiction	No	311(72.3%)	119(27.7%)	430(100%)
	Yes	9461(71.7%)	3737(28.3%)	13198(100%)

The above Table 4.1 shows that of total respondents 7464 are male and 6164 are female respondents included in our study. Of total male respondents 30.6% are unemployed and 69.4% are employed. 25.5% of female respondents are unemployment and the rest 74.5% of female are employed.

Drug addiction is another factor with regard to the determinants of unemployment status in Ethiopia .From the above table4.1 we can observe that 28.3%of drug addiction has employed and approximately 71.7%is unemployed Whereas27.7% of unemployed and72.3 of employed are not drug addiction.

Table 4.1 also reveals that the unemployment status varies by their educational status. The highest percentage unemployment was observed who have illiterate (85%) and the lowest percentage of unemployment was observed who's who have higher education (22.7%).

The above table 4.1 reveals that unemployment status differs by marital status

For instance, 22.4% of single respondents had unemployment, 34.2% of Married and 36.7% widowed respondents had unemployment.

The above table 4.1 reveals that unemployment status differs by Region The highest percentage unemployment was observed in Afar 29.9% and the lowest percentage of unemployment was observed inBenishangulgumze 22.6%.

4.2. Inferential statistics

Inferential statistics is statistical method that deals with making inference or conclusion about the population based on data obtained from a limited number of observations that come from the population.

Table 4.2.1Chi-square test of independence

Variable	Pearson chi-squar	Df	p-value
Region	15.630^{a}	10	0.0011
Sex	.023a	1	0.878
Education	2.3502a	4	0.0
Marital status	0.486E2a	2	0.678
Durg addiction	1.084a	1	0.002

Hypothesis test:

$H_{0:}$ there is no significant association between response and explanatory variables.

H_1: there is a significant association between response and explanatory variables.

From table 4.2.variables which had p-value of Pearson chi-square less than α=0.05 was considered to have significant association with unemployment status, where as variables which had p-value of Pearson chi-square greater than a value of α=0.05 had no significant association with unemployment status. Since region, age, drug addiction and education level has a significant association with unemployment status. However sex andMarital has no significant association with unemployment status at 5% level of significance.

4.2.1. Binary Logistic Regression

Logistic regression examines the relationship between one or more predictor variables and a binary response variable and hence the logistic regression equation can be used to examine how the probability of an event changes as the predictor variable changes.

Table 4.2.2**Dependent Variable Encoding**

Original Value	Internal Value
Employed	0
Unemployed	1

From the above table 4.2.2dependent variables shows that urban youth unemployment coded

As: - 0=employed and 1 = unemployed

Table 4.2.3 Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	13628	100.0
	Missing Cases	0	.0
	Total	13628	100.0
Unselected Cases		0	.0
Total		13628	100.0

a. If weight is in effect, see classification table for the total number of cases.

Table 4.2.3 provides a basic descriptive window showing how many subjects Analyzed, and how many subjects were missing. As can be seen above, all 13,628 subjects in this analysis were included, with zero subjects missing, and none unselected.

Table 4.2.4 Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 4[a]	Step	-1.902	1	.168
	Block	187.676	15	.000
	Model	187.676	15	.000

a. A negative Chi-squares value indicates that the Chi-squares value has decreased from the previous step.

Ho: βj=0 (the independent variables could not predict urban youth unemployment)
H1: at least one independent variable can predict urban youth unemployment.

From Table 4.2.4 we have added one new variable to the model, which has reduced the -2log likelihood by-1.902with 1 degree of freedom. The p-value for the result of adding to the model is depicted in the table4.8 and we can see that this is 0.168 which is greater than the α-level of significance (0.05).hence we would conclude that, the addition of each independent variable to the model is not statistically significant. In other words, this variable does not explain variation in the urban youthunemployment. And a negative chi-square value indicates that the chi-square value has decreased from the previous step

4.2.3 Model adequacy checking

After a logistic regression model has been fitted, a global test of goodness of fit of the resulting model should be performed. It is necessary to see the appropriateness, adequacy and usefulness of the fitted model. The most commonly used technique are the likely hood ratio test and the Hosmer-Lemeshow test.

4.2.3.1 Likelihood Ratio Test

Table4.2.5 Model Summary of Logistic Regression

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
4	4144.674[b]	.014	.050

b. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

The Cox and Snell R^2 or Nagelkerks R^2 is analogous statistic in logistic regression to the coefficient of determination R^2 in linear regression. From Table 4.2.5 Cox and Snell R^2 indicate that, about 1.4 % of variable in the dependent youth unemployment is explained by the independent variable. Also nagelkerke R^2 indicate that 5.0%.

4.2.3.2Hosmer and Lemeshow Goodness of fit test

The Hosmer-Leme show test is performed by dividing the predicted probabilities into deciles groups based on percentile ranks and then computing a Pearson chi-square that compares the predicted to the observed frequencies.

Table 4.2.6 Hosmer and Leme show test of Logistic Regression

Step	Chi-square	df	Sig.
4	12.497	8	1.30

Table 4.2.6

The hypothesis to be tested as follows:

Ho=the model is good fit

H1=the model is not a good fit

Since the p-value (1.30) is greater than α=0.05 the statistical level of significance and it implies that fail to reject the null hypothesis (Ho). So we conclude that the model is a good fitted model

The "Hosmer and Lemeshow Test" is a measure of fit which evaluates the goodness of fit between predicted and observed probabilities in classifying the response variable. Similar to the - 2 log likelihood test, we want this chi - squared value ($X^2_{8,0.05}$=12.497) to be low and non - statistically significant (p- value=1.32) if the predicted and observed probabilities match up nicely.

In this case we see that the test is statistically insignificant (p >.05), suggesting that the probabilities of predicted versus observed values of the response variable match up as nicely as we would like. Therefore,our fitted logistic regressionmodel is good fit.

Table 4.2.7Result of Binary logistic regression

Estimates, Standard error, Wald, degree of freedom, P-value, estimated odds ratio

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 4[a]	Addis Ababa (ref)			44.309	10	.000	
	Afar	-.226	.243	.865	1	.352	.797

Amahar	-1.267	.391	10.508	1	.001	.282
Ormoiya	-.363	.218	2.764	1	.096	.696
Somali	-.552	.215	6.571	1	.010	.576
Benishangul-Gumuz	-1.973	.491	16.123	1	.000	.139
SNNPR	-1.063	.365	8.494	1	.004	.346
Gambela	-.769	.246	9.765	1	.002	.463
Harari	-1.119	.364	9.447	1	.002	.327
tgriye	-1.373	.430	10.209	1	.001	.253
Deridewa	-.552	.231	5.702	1	.017	.576
Age	-.095	.012	64.135	1	.000	.910
Primary(ref)			30.973	4	.000	
secondary	-.644	.121	28.264	1	.000	0.525
Preparatory	-.494	.126	15.459	1	.000	0.610
Adult education	.679	.240	7.976	1	.005	1.971
Higher education	-.463	1.015	.208	1	.648	.630
Drug addiction	.347	.109	8.463	1	.004	1.415
Constant	-.691	.374	3.412	1	.065	.501

Variable(s) entered on step 1: ID101, UE204, UE205, education, marital, drug.

b. Variable(s) entered on step 5: durg.

c. Stepwise procedure stopped because removing the least significant variable result in a previously fitted model.

Table4.2.7 contains the estimated coefficients (under the column heading β̂) and estimated values of the logistic regression model that predict urban youth unemployment. The standard error of the estimates (under the column heading S.E) will help in computing the Wald Statistics. The Wald statistic, which is the square of the ratio of the coefficient to its Standard error, has a chi-square distribution with one degree of freedom.

The significance of the Wald statistic (under the column labelled Sig) tells the importance of the predictor variable in the model. The column exp (β̂), is the factor by which the odds of experience of sexual violence change when the ith independent variable increases by one unit. If $\tilde{\beta i}$ is positive, exp ($\tilde{\beta i}$) will be greater than one, which means the odds of experience of urban youth unemployment increases. If βi is negative, exp ($\tilde{\beta i}$) will be less than one, which means the odds of experience of urban youth unemployment decreases.

The fitted model can be written as:-

Logit (π) = $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots + BpXp$

Where: $X_1, X_2 \ldots\ldots X_P$ are the independent variable

$\beta_1, \beta^2 \ldots \beta P$ are coefficient of independent variable and

βo is constant (intercept).

Logit (π) = $-0.691+-12.67X_{12} + -.552\ X_{14}+ -1.973X_{15}+-1.063\ X_{16}+-.769\ X_{17}-1.119X_{18}+ -1.373\ X_{19} +-.552x_{10} + -.095X_{21} +.644X_{31} +.494X_{32} +.679X_{33}+.374X_{41}$

Where βo=-0.691

X1= Region: X12, X14,X15,X16,X17,X18,X19,X10 =Amahar ,Somali, Benishangulgumze ,SNNPR, Gambela, Harari ,AddisAbeba and Deridewa respectively

X21=Age

X3=Education: - X31, X32and X33= secondary, preparatory and Adult education respectively,X_{41} = drug addiction

4.2.2. Interpretation for coefficients of significant variables

Agehad showed negative effect on urban youth unemployment Thenegative coefficient indicates that a one unit decrease inage will decrease theurban youth unemployment and other factors being constant. The result also implied that, a unitdecrease in the age of youth (mane age around the beginning of youth age) caused an decrease by.095of youth unemployment Amahar, Somali, Benishangul-Gumuz, SNNPR, Gambela, Harari and tgriye has negatively affectyouth unemployment decreased by1.267, 0.552, 1.973, 1.063, 0.769, 1.119, 1.373and 0.552 respectively ascumbared to reference category (addis Ababa)

Secondary and preparatory has negatively affect youth unemployment decreased by0.644and0.494 respectively as cumbared to reference category (addisabeba) WhereasAdult education has positively affect youth unemployment increased by 0.679as cumbared to reference category (addisabeba)

Drug addiction has positively affecton youth unemployment .The positive coefficient indicates that a one unitincrease indrug addiction youth will increase the urban youth unemployment by 0.347 and other factors being constant.

Age had showed negative effect on urban youth unemployment with significance level at 5%The negative coefficient indicates that a one unit decrease in age will decrease the urban youth unemployment and other factors being constant. The result also implied that, a unit decrease in the age of youth (mane age around the beginning of youth age) caused an decrease **by**.095 of youth unemployment

4.2.2.1. Wald test

Since Wald test is used to test the statistical significance of individual coefficient (β) in the model and the test statistic is a chi-square statistic.

To conclude that the given coefficient is significant to model based on the following:-

i. The chi-square (Wald) statistics must be greater than tabulated statistic($\chi^2_{0.05,1}$)

ii. P- Values of coefficients are less than the level of significance, α=0.05.

For β_1:- From the parameter estimation above; the chi-square statistics (Wald) =44.309 is greater than $\chi^2_{0.05,1} = 3.84$ the p-values for the β_1 =0.000 is less than 0.05 level of significance. Thus based on this result we see that the coefficient of region is significant to the model.

For β_2: From the parameter estimation above; the chi-square statistics (Wald) = 64.35 is greater than $\chi^2_{0.05,1} = 3.84$ the p-values for the β_2 =0.000 is less than 0.05 level of significance. Thus based on this result we see that the coefficient of age is significant to the model.

For β_3: From the parameter estimation above; the chi-square statistics (Wald) = 30.973 is greater than $\chi^2_{0.05,1} = 3.84$ the p-values for the β_3 =0.000 is less than 0.05 level of significance .Thus based on this result we see that the coefficient of education level is significant to the model.

Forβ_4 : From the parameter estimation above; the chi-square statistics (Wald) = 8.463is greater than $\chi^2_{0.05,1} = 3.84$ the p-values for the β_3 =0.004 is less than 0.05 level of significance .Thus based on this result we see that the coefficient of drug addiction is significant to the model.

4.2.2.2 The Odds Ratio Interpretation:-

We can interpret the odds ratio of region obtained from the above table using the reference category Tigray. Urbane youth unemployment who followsAmaharare0.282times less likely than that of followed addisabebe(the reference one) controlling for other variable in the model.

The estimated odds ratio Somali are 0.576 times less likely than that of followed DireDewa (the reference one) controlling for other variable in the model, The estimated odds ratioof Benishangulgumzeare 0.139times less likely than that of followed addisabeba(the reference one)The estimated odds ratio of SNNPR are 0.346times less likely than that of followed addisabebe(the reference)The estimated odds ratio of Gambela are 0.463times less likely than that of followed addisabebe (the reference) The estimated odds ratio of Harari are 0.327times less likely than that of followed addisabebe(the reference) The estimated odds ratio of Tigray are 0.253times less likely than that of followed addisabebe (the reference)The estimated odds ratio of Deridewa are 0.576times less likely than that of followed addisabebe (the reference)

We can interpret the odds ratio of education level obtained from the above table using the reference category Primary (ref) Urbane youth unemployment who follow secondary are 1.904 preparatory times more likely than that of followed Primary (ref) controlling for other variable in the model, Urbane youth unemployment who follow preparatory are 1.639 times more likely than that of followed Primary (ref) controlling for other variable in the model Urbane youth unemploymentthatfollows Adult educationis1.971 times more likely than that of followed Primary (ref) controlling for other variable in the model

4.3. Model adequacy checking

After a logistic regression model has been fitted, a global test of goodness of fit of the resulting model should be performed. It is necessary to see the appropriateness, adequacy and usefulness of the fitted model. The most commonly used technique are the likely hood ratio test and the Hosmer-Lemeshow test.

4.3.1. Likelihood Ratio Test

Table4.3.1. Model Summary of Logistic Regression

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
4	4144.674[b]	.014	.050

b. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

The Cox and Snell R^2 or Nagelkerks R^2 is analogous statistic in logistic regression to the coefficient of determination R^2 in linear regression. From Table 4.9 Cox and Snell R^2 indicate that, about 1.4 % of variable in the dependent youth unemployment is explained by the independent variable. Also Nagelkerke R^2 indicate that 5.0%.

4.4Discussion

This study is undertaken to determine the determinant of urbane youth unemployment in Ethiopia. Data from the CNA 2018 are used for analysis. A total of 13121 youth aged 18-32. For these purpose, descriptive statistics and binary logistic regression are used. The results obtainedare discussed as follows.The finding of this study shows that sex had a notsignifican related to Urban youth unemployment. This finding opposedto the result of a studyin DebreBirhan town by[6], where they showed thatsex, , education, job preferences and access to business advisory services significantly determine youth unemployment

The finding of this study shows that marital status had insignificantly related to urban youth unemploymentThis finding corresponds with the result of a study in Addis Ababa by [22] where they showed thatmarital status were found insignificantly related to youth unemployment

The finding of this study shows that region age and education level had significance effect with youth unemployment.by [16],where they showed thatconducted the binary logistic regression to assess the determinants of youth unemployment at Ambo, Ethiopia. Their result showed that among the demographic variables, age of the respondents and migration status were significantly related to youth unemployment whereas marital status of the respondents was not significant. From the human capital variables included in the model, education and health status of the respondents were significantly related to youth unemployment,

CHAPTER FIVE

5. Conclusions and Recommendations

5.1 Conclusions

The main objective of this study was to investigate the factors that determine unemployment status in Ethiopia. We conclude this study was an attempt to examine the impact of some determinants that determine unemployment status in Ethiopia . In this analysis wehave looked at logistic regression models that can be applied when our outcome is represented by a binary variable. We see in the above interpretation the Proportional odds assumption is justified binary regression models can be a powerful means of summarizing relationships that utilizes all the information present in the binary outcome. Furthermore, the findings indicate that the employment status is significant associated with region,age,drug addiction and education level, occupational status and drug addiction. However sex and marital status has no significant association with employment status.

5.2. RECOMMENDATION

The researcher would like to recommend the Ethiopian urban to decrease unemployment based on the finding.

- The government should give more support and emphasis on those Ethiopian urban with high rate of unemployment. Additionally, further research on socio-cultural practices, distribution of education, and other related factors should be emphasized. In order to decreases youth unemployment levels in this urban area.

- The government should increase and improving the level of education will create youth employment and better chances to reduce youth unemployment. Beside, creation of new jobs can be achieved through education.

- The government should create awareness for drug addiction how affecte youth employment (cause for youth unemployment) and give a successive advice for youth to reduce their drug addiction.

REFERENCES

1. Agresti, A.1990.Categorical Data Analysis. New York: John Wiley andSons.
2. Alan Agresti, A.1996. An Introduction to Categorical DataAnalysis. New York:John Wiley and Sons.
3. AlemnewGetnet (2014), Socio-economic & Demographic Determinants of Graduate youths' unemployment: In the case of DebreMarkos town, AmharaRegionalState, University of Gonder.
4. Alhawarin, I., &Kreishan, F. (2010), An analysis of long-term unemployment (ltu) in Jordan's labor market. European Journal of Social Sciences, 15(1), 56-65.
5. AmanuelDisassaAbshoko (2016), Determinants Of Youth Unemployment; Evidence From Ethiopia,
6. Asalfew A. (2011), Demographic and socio-economic determinants of youth unemployment in DebereBirhan town, Master thesis, Addis Ababa University, Addis Ababa.
7. Bakare A.S (2011). The Determinants of Urban Unemployment Crisis inNigeria:An Econometric Analysis Department of Economics, AdekunleAjasin University, Nigeria.
8. BerhanuDenu, Abraham Tekeste and Hannah Van DerDeij (2005). Characteristics and Determinants of Youth Unemployment, Underemployment, Underemployment and Inadequate Employment in Ethiopia, Employment Policies Unit Employment Strategy Department.
9. Bluman, 3rd Edition
10. Burnham and Anderson 1998.
11. Dejene T., J.PaulMansingh and Warkaw L., Determinants of Youth Unemployment: The Case of Ambo Town, Ethiopia. International Journal of Economics and Business Management, 2016, 2(2), 162-169.
12. Echebiri, R.N. (2005). Characteristics and Determinants of Urban Youth Unemployment in muahia, Nigeria: Implications for Rural Development and Alternative Labour Market Variables. OkparaUniversity of Agriculture, Umuahia, Abia State
13. Getinet Haile. 2003. The Incidence of Youth Unemployment in Urban Ethiopia. Paper presented at the 2nd EAF International Symposium on Contemporary Development Issues in Ethiopia, 11-13 July, Addis Ababa, Ethiopia.
14. Hosmer, W.D. and S. Lemeshow (2000). Applied Logistic Regression. 2nd Ed, John Wiley and Sons, New York
15. ILO (1992). Surveysof Economically Active Population. Employment, Unemployment and Underemployment: ILO manual on Concepts and Methodology. Geneva.
16. ILO (2010). Global Employment Trends for Youth. Special Issue on the Impact of the Global Economic Crisis on Youth, Geneva.

17. Labor Force Survey (2013), Statistical Report on the 2013 National Labour Force Survey, Central Statistical Agency, Addis Ababa, Ethiopia
18. Ministry of Youth, Sports and Culture (2004), National Youth Policy, Addis Ababa, Ethiopia.
19. Muhammad ShahidMaqbool, TahirMahmood Abdul Sattar And M.N.Bhalli(2013) ,Determinants Of Unemployment Empirical Evidences From Pakistan, Pakistan Economic and Social Review, Volume 51, No. 2 (Winter 2013), pp.191-207
20. Nzinga H. Broussara and TsegayGebrekidanTekleselassie (2012), Youth Unemployment: Ethiopia Country Study, International Growth Center
21. Shumet, G. (2011). A Glimpse of Urban Youth Unemployment in Ethiopia. Ethiopian Journal of Development Research, vol.33, No.2.
22. TegegnGebeyaw (2011), Socio-Demographic Determinants ofurbanUnemployment: The Case of Addis Ababa Ethiopian Journalof Development Research Vol.33, No.2, October 2011
23. Toit, R. (2003).Unemployed Youth in South Africa: The Distressed Generation, Paper presented at the MinnesotaInternational CounselingInstitute, Minnesota.
24. WB. (2007), urban labor markets in Ethiopia: Challenges and prospects: The World Bank.

Appendix

Variables not in the Equation

			Score	df	Sig.
Step 2[a]	Variables	formai(1)	7.432	1	.006
	Overall Statistics		7.432	1	.006
Step 3[b]	Variables	marital	.155	2	.925
		marital(1)	.154	1	.695
		marital(2)	.112	1	.738
		formai(1)	7.392	1	.007
	Overall Statistics		7.591	3	.055
Step 4[c]	Variables	UE204(1)	1.898	1	.168
		marital	.176	2	.916
		marital(1)	.081	1	.776
		marital(2)	.013	1	.909
		formai(1)	7.636	1	.006
	Overall Statistics		9.259	4	.055
Step 5[c]	Variables	UE204(1)	1.424	1	.233
		marital	.109	2	.947
		marital(1)	.103	1	.749
		marital(2)	.064	1	.801
	Overall Statistics		1.621	3	.655

a. Variable(s) removed on step 2: formai.

b. Variable(s) removed on step 3: marital.

c. Variable(s) removed on step 4: UE204.

Categorical Variables Codings

		Frequency	Parameter coding					
			(1)	(2)	(3)	(4)	(5)	(6)
Region	Tigray	967	1.000	.000	.000	.000	.000	
	Afar	488	.000	1.000	.000	.000	.000	
	Amhara	2650	.000	.000	1.000	.000	.000	
	Oromo	3619	.000	.000	.000	1.000	.000	
	Somalie	551	.000	.000	.000	.000	1.000	
	Benishangul-Gumuz	461	.000	.000	.000	.000	.000	1
	SNNPR	1599	.000	.000	.000	.000	.000	
	Gambela	525	.000	.000	.000	.000	.000	
	Harari	438	.000	.000	.000	.000	.000	
	Addis Ababa	1898	.000	.000	.000	.000	.000	
	Dire Dawa	432	.000	.000	.000	.000	.000	
levelofeducation	primary school	4025	1.000	.000	.000	.000		
	secondary school	3694	.000	1.000	.000	.000		
	Preparatory school	437	.000	.000	1.000	.000		
	illiterate	68	.000	.000	.000	1.000		
	higherleveleeducation	5404	.000	.000	.000	.000		
Marital Status	married	5973	1.000	.000				
	single	6969	.000	1.000				

	widowed	686	.000	.000				
Sex	Male	7464	1.000					
	Female	6164	.000					
Drug addiction	yes	430	1.000					
	n0	13198	.000					