



ADDIS ABABA UNIVERSITY
GRADUATE STUDIES PROGRAM

***DEMOGRAPHIC AND SOCIO-ECONOMIC PREDICTORS
OF ATTITUDINAL CHANGE ABOUT CONDOM USE IN
FOUR REGIONS OF ETHIOPIA***

By
Yonas Kassa

***A thesis submitted to the Graduate Studies Program of Addis
Ababa University in partial fulfillment of the requirements for
the Degree of Masters of Science in Statistics***

Addis Ababa

September 2007.

Addis Ababa University
Graduate Studies Program

Title of the thesis Demographic and socio-economic predictors of
attitudinal change about condom use in four regions
of Ethiopia

Name of the candidate Yonas Kassa

Department Statistics

Faculty Science

Approve by the Board of Examiners:

Department Head

Signature

Internal Examiner

Signature

External Examiner

Signature

Tel. 251-1-239465

Ababa Addis, Ethiopia

P.O.Box 1176

ACKNOWLEDGEMENT

I wish to express my sincere thanks to my advisor, Pro. Eshetu Wencheko, for his many, constructive comments and suggestions throughout the study.

Special thanks also due to Ato Gemechiss Tefferi, who has facilitated the opportunity to get data from CSA.

Finally, I wish to express my heartfelt gratitude to all my friends and my family for their continuous support and encouragement.

Table of Contents

	Page
ACKNOWLEDGEMENT	i
ABSTRACT	ii
CHAPTER I	
Introduction	1
1.1 Background of the study	1
1.2 Objective of the study	3
1.3 Limitation of the study	4
1.4 Definition of Terms	4
CHAPTER II	
Review of the literature	5
CHAPTER III	
The data and methodology	11
3.1 The Data	11
3.1.1 Source of the data	11
3.2 Variables in the Study	12
3.2.1 Dependent Variable	12
3.2.2 Independent Variables	13
3.2.3 Interaction	15
3.3 Methodology	15
3.4 Logistic Regression	15

3.5 Parameter Estimation and Statistical Inference	16
3.6 Model Selection	18
3.7 Model Checking	20
CHAPTER IV	
Statistical data analysis	24
4.1 Introduction	24
4.2 Variables Selection	24
4.2.1 Bivariate Findings	24
4.2.2 Multivariate Findings	26
4.3 Model checking and diagnostics	30
CHAPTER V	
Discussion, conclusion and recommendation	34
5.1 Discussion	34
5.2 Conclusion	35
5.3 Recommendation	35
Reference	37
Annex	
Table A Correlation matrix	39

Table B Model Summary	41
Table C Hosmer and Lemeshow Test	41
Table D Contingency Table for Hosmer and Lemeshow Test	41
Table E. Classification Table	43
Table F. Variables in the Equation	44
Table G Model if Term Removed	44
Observed Group and Predicted Probabilities	46

ABSTRACT

In an attempt to identify Demographic and Socio-economic predictors of awareness about condom use in four Regions of Ethiopia, this study focused on a sample of legislators, professionals and unemployed individuals in the age group 15-59 years. In this study data gathered by the Central Statistical Agency (CSA) for 2004 Behavioral Surveillance Survey were used. From the analysis of the data it was found that the most important demographic and socio-economic predictors which determine awareness about the use of condom in the four regions considered are educational level, marital status, awareness about the availability of condom, age group, sex, region and use of other contraceptive method. In order to increase the level of awareness about condom use strategies and policies must be developed which focus on educating people, increasing distribution of condom, and encouraging governmental/non governmental organization to participate in the protection of HIV/AIDS and other sexually transmitted diseases by using condom.

CHAPTER I

INTRODUCTION

1.1 Background of the study

AIDS is an extraordinary kind of crisis; it is both an emergency and a long-term development issue. Despite increased funding, political commitment and progress in expanding access to HIV treatment, the AIDS epidemic continues to outpace the global response. No region of the world has been spared. The epidemic remains extremely dynamic, is expanding fast and also changing its character as the virus exploits new opportunities for transmission.

Since the first AIDS diagnosis in 1981, about 20 million people died and almost 38 million people (range 34.6 – 42.3 million) are living with HIV (MOH, 2005).

Rates of infection are still on the rise in many countries. All parts of the world are not equally infected by this pandemic. There is a considerable regional variation in its distribution. Sub-Saharan Africa is the region where the highest number of victims of HIV/AIDS is found. Although the region accounts only for 10% of the world population, it comprises almost 25.8 million people of the victims of HIV/AIDS in the world. In 2005 an estimated 3.2 million people in the region became newly infected, while 2.4 million died of AIDS. Among the younger generation (15-24 years) the percentage of HIV infected women and men account for 4.6% and 1.7%, respectively (UNAIDS, 2005).

Ethiopia has one of the highest HIV prevalence's in sub-Saharan Africa. In Ethiopia, the highest prevalence of HIV (12.1%) is observed among youth whose age belongs to 15-24 years (Family Health International (FHI), 2004). According to the UNAIDS 2004 report, Ethiopia stands fifth, just next to South Africa, Nigeria, Zimbabwe and Tanzania where the highest victims of the disease are found. The 2003 National Sentinel Surveillance (NSS) results document of the Ministry of Health (MOH) report on "AIDS in Ethiopia"

indicated that the number of people living with HIV/AIDS was about 1.5 million out of which about 96,000 are children whose age is below 15 years. The estimated number of new HIV/AIDS cases among the adult population in 2003 was 98,000 (46% male and 54% female) while it is about 25,000 among children. Some 245,000 People Living with HIV/AIDS (PLWHA) were in need of Antiretroviral Treatment (ART) and some 90,000 adults and children died of AIDS in 2003. There were also an estimated 593,000 AIDS orphans in the same year (MOH, 2004:11). Based on HIV data collected at antenatal clinics, national adult HIV prevalence in Ethiopia in 2005, the infection level was found to be more than five times higher in urban (10.5%) than in rural (1.9%) areas (MOH, 2004).

Condom is the basic means for protecting HIV/AIDS and other sexually transmitted diseases (STDs). Condom is not only a protection against STDs but also unwanted pregnancy. The need for condom is growing as HIV/AIDS and other STDs spread; and this is also true for a country like Ethiopia where heterosexual relation is the primary mode of spread for HIV. Making condom more accessible, lowering the cost of condom, promoting awareness extensively and helping to overcome social and personal obstacles that limit the use of condom, are some of the widely advocated strategies that can be used to save life loss due to HIV/AIDS.

The Ethiopian Social Marketing Project (ESMP), implemented by Dioptic Laboratories (DKT) International in collaboration with Population Services International (PSI) has emerged as one of the most successful social marketing projects in Africa since 1990. During the 33 months of collaboration with AIDSCAP, the ESMP sold close to 42 million condoms through more than 10,000 outlets in 533 different cities, towns and villages. In the last full year of project activities, the ESMP sold over 15 million condoms, averaging 1.25 million per month, and 25 percent more than the target of one million per month. Condom social marketing in Ethiopia overcame serious constraints to implementation including the lack of a developed commercial marketing and distribution system, the expanding poverty, religions and relative unfamiliarity with condoms on the part of target populations. Of particular significance is the fact that sales through bars and

hotels accounted for nearly 25 percent of all ESMP sales and also small number of users relative to the whole population. AIDSCAP support assisted DKT/PSI to secure longer term funding from USAID and the Dutch Government for the continued operation and expansion of social marketing in Ethiopia in 1999. Just over a million condoms were also distributed free of charge through public and private outlets (FHI/AIDSCAP program in Ethiopia, 2000).

In general, the overall condition of the use of Condom in Ethiopia calls for scaled up and coordinated activities to improve the service. It is, therefore, crucial to have data on knowledge, attitude, practice and other issues related to the use of condom among the study group. Hence, the present study is initiated with the main objective of investigating demographic and socio economic factors on condom use among the study group. The findings of the study could be useful in designing intervention programs that will promote condom in the fight against HIV/AIDS and other Sexually Transmitted Diseases (STDS) and also unwanted pregnancy.

This study has five chapters. Brief discussions on the revised studies are given in chapter two. Discussion about the data for the analysis, factors expected to predict the awareness level of the use of condom and the methodology applied for the study are covered in chapter three. In chapter four, we give results of the data analysis, a goodness of fit test of a model together with its adequacy and a thorough discussion on the results will lead us to the end of the chapter. Conclusion and Recommendation of the study are presented in chapter five. Tables and Graph are appended.

1.2 Objective of the study

The overall aim of the study is to investigate demographic and socioeconomic factors on condom use.

More specifically, the study attempts to:

- Investigate the awareness level of the study group about condom use

- Identify the degree of the relationships among the factors under investigation
- Select the factors which have the most significant influence on condom use
- Provide necessary suggestions and recommendations

1.3 Limitation of the study

This study has a limitation in the sense that the target population consists of only legislators, professionals and the unemployed in the age group 15-59 years.

1.4 Definition of Terms

Sexually transmitted diseases (STD) are diseases that can be transmitted sexually, like Syphilis, Hepatitis B, Chlamydia and HIV/AIDS infection and others.

Contraceptive method is the method protecting from unwanted pregnancy, for instance pills, condom, and other methods.

CHAPTER II

REVIEW OF THE LITERATURE

A study by Muhwava (2004) on Condom Use within Marriage and Consensual Unions in Zimbabwe, showed that the differentials by sex are significant, showing females are less likely to use condoms always or occasionally with their regular partners compared to males. The study indicated that the differentials by place of residence are also significant, showing that the respondents in rural areas are less likely to use condoms always or occasionally with their regular partners compared to those respondents in urban areas. It was also found that respondents with high secondary and tertiary education are more likely to use condoms always or occasionally with their spouse or cohabiting partners than those with no or primary education. Condom use with regular partner tends to increase as the level of education increases. Those working in agriculture or subsistence farmers are less likely to have used condoms with their spouses or regular partners than those in other occupational categories. Couples in professional employment, sales and service workers are significantly likely to use condoms always or occasionally with their regular partners or spouses compared to agricultural workers. Christians are more likely to have ever used condoms with their spouses or regular partners relative to non-religious traditionalists.

Pullum, et al (2004) conducted a study of Factors affecting Condom use in Primary Sexual Relationships in southern and eastern Africa. This study used data collected in Kenya, South Africa, Tanzania, Uganda, Zambia and Zimbabwe. From the multivariate analysis they found that reported levels of condom use tend to be higher in urban areas than in rural areas, and higher for males than for females. Apart from sampling variation, they would expect men's and women's levels of current use and of ever-use with the primary partner to be fairly close. The sex differential, which is most pronounced for current use, may be due to over-reporting by men, but it is also possible that some women

under-report it because it is a male method. They could also investigate a favorable attitude toward condoms is associated with a subjective risk of HIV infection and a positive attitude toward family planning.

A logistic regression analysis by Juarez and Martin (2006) showed that educational attainment, degree of knowledge of HIV transmission and prevention, and condom use at first sexual intercourse were found to be significantly associated with current condom use. Regarding the relationship context, the analysis revealed that adolescent males in steady relationships were less likely to use condoms, less likely to regard themselves at risk of HIV infection, and more concerned about pregnancy prevention than adolescents in casual relationships. Differentials in condom use by type of relationship, however, did not result from a higher rejection of contraception by steady partners but from their higher likelihood to rely on other contraceptive methods.

Arjun and Ipshita (2006) studied the effect of AIDS awareness on condom use intention among truck drivers in India. They found that beliefs and feelings mediate the effect of AIDS awareness on condom use intention. Although there was a significant relationship between the frequency of sexual episodes and perceived vulnerability, their effects on condom use intention were non-significant.

Mnyika, et al (1997) studied on determinants of high-risk sexual behavior and condom use among adults in the Arusha region, Tanzania. They could observe that significantly more men (49%) than women (25.2%) reported having multiple sexual partners. Urban men were reported to have multiple sexual partners than men in rural areas. In both men and women, early sexual debut was associated with having multiple sexual partners while traveling, alcohol use, and sex under the influence of alcohol were significantly associated with multiple sexual partners in men only. Of the 1551 respondents, 320 (20.6%) reported having ever used a condom and of the 320 respondents who had ever used a condom, 34 (10.6%) reported having used it at the last sexual intercourse. Significantly more men (34.1%) than women (14.1%) reported having ever used a condom. In both men and women having multiple sexual partners is associated with

increased condom use. For both men and women, frequent discussion of AIDS with family members or friends was associated with increased condom use.

Mannin, et al (2000) showed that approximately 31% of respondents used no contraceptive method at first intercourse. Roughly half (52%) of adolescents who had just met their sexual partner used no method, compared with 24% of those who were going steady. Whereas 75% of teenagers who practiced contraception at first intercourse used condom, 17% relied on the pill. Individual-level factors that influenced contraceptive use at first intercourse were age at first intercourse, race or ethnicity, family type, parents' education, grades in school and receipt of birth control education prior to first intercourse. Differences in age and race or ethnicity of respondents were not significantly related to method use at first intercourse. One exception was that adolescents who first had sex with a man six or more years older had reduced odds of practicing contraception. Variables associated with pill use rather than condom use were age at first sex, race, family type, mother's education and school grades.

Manchikanti (2005) found a strong relationship between the experience of sexual violence and adverse reproductive health outcomes among young women and adolescents in Colombia after controlling for socio demographic characteristics and age at first sex. Young women who experienced sexual violence differed significantly from those who did not in terms of unwanted pregnancy, condom use at last sex, and recent STI diagnosis and current use of modern contraception.

Brandy and Laurie (2002) reviewed eighteen studies about the relationship between religious denomination and sexual behavior. The most consistent findings from these studies indicate that black and white Catholic or fundamental Protestant females initiate sexual intercourse later than their counterparts with another denomination or who do not affiliate with any religion. Their Findings also show Protestant females are less likely to use contraception when they do engage in sexual activity, and for Catholics, these findings are mixed. The findings to date for male sexual behavior are too inconclusive to make generalizations. The review indicated that more frequent attendance at religious

services is associated with more conservative sexual attitudes and less frequent intercourse, later initiation of intercourse for white males and for females across racial/ethnic groups, and decreased contraceptive use for females and increased use for males. The importance of religion in the adolescents' lives is inconclusive. Two studies showed an association between greater importance of religion and a delay in first intercourse. One of these two studies showed that boys who said religion was very important were more likely than others to have had recent sexual activity, and while the other study failed to show any association between the importance of religion and condom use.

Mumtaz, et al (2004) explored the influence of gender access to resources and couple-level dynamics in Zimbabwe and Uganda. The study showed that gender inequalities in access to resources and couple dynamics influence the use of condoms among couples in Zimbabwe and Uganda.

Looking at the Ethiopian expression, according to Taffa et al (2002), a third of 561 in- and out-of-school youth (15–24 years) in Addis Ababa who completed a questionnaire reported that they had sexual intercourse in the past. Half of the sexually active used condoms during recent intercourse. This study showed that being out-of-school, male, age group 20–24 years, alcohol use and khat (amphetamine-like substance) consumption predicted the likelihood of engagement in sexual activity. Males reported frequent condom use during recent sexual intercourse. Self-efficacy, skills, and barriers predicted 23% of the variance in intentions to use condoms. Self-efficacy was also associated with past condom use. Psychosocial constructs predicted more variations in condom use intention for males than for females. In general, self-efficacy was found to be the strongest predictor of the constructs, whereas attitude and social influences were the weakest.

Belachew Gebre (2002) conducted a community based cross-sectional descriptive study with the aim of narrowing the existing knowledge gap in the area of condom accessibility and utilization among people residing in urban, semi-urban and rural areas of Ethiopia.

The study revealed major differences in level of awareness about HIV/AIDS, the risk factors for HIV infection and perceived vulnerability, between rural and urban residents and married and non-married people. Almost in all cases, those who live in rural areas were on disadvantage. Nearly 90% of those who know about AIDS also know about condoms, radio being the main source of information in all the three study areas. About 76% of those who have heard about condoms have also seen condoms. Nearly 72% of those people who have sexual partner(s) and used condom reported to have discussed condoms with their partners. A third of the respondents used condom at least once in their lifetime of which 59% used condoms during their recent sexual encounter, and 97% used to protect themselves from STD/HIV/AIDS. The proportion of condom users is higher among people with high school and college level education. Nearly half the females believe that males are not willing to use condoms and about third pointed out that male make the final decision to use or not to use condom. Of the people who would not recommend condoms to others, about a quarter believe that condoms motivate people to be promiscuous.

Jegu Degu (2002) reported that the level of knowledge of commercial sex workers about HIV/AIDS (97.2%) and condom (96.2%) was stated to be very high compared to their experience of using condom (32.5%) in general and frequent use of condom (12.8%), in particular. Condoms were reported to be readily available at a cheap price. Among the parents of commercial sex workers who did not like to use condoms, the majority (81%) were farmers who come to those small towns from the surrounding rural areas. A higher condom use rate was observed as the level of educational status of commercial sex workers increased from lower to higher grades.

Mesfin Samuel, et al (2004) found higher levels of positive attitudes toward and self-efficacy for condom use were associated with higher levels of both past and intended condom use. Knowledge of individuals who live with or have died of HIV/AIDS was also directly related to higher levels of intended condom use. In contrast, involvement in risky sexual behaviors and use of other contraceptive methods predicted higher levels of past condom use. Effects of reproductive health understanding and family socioeconomic

status on past and intended condom use were indirect through their positive influence on psychosocial factors, particularly self-efficacy and attitudes. Overall psychosocial and contextual model accounted for 40 and 62 percent of the variance in past and intended use, respectively.

CHAPTER THREE

THE DATA AND METHODOLOGY

3.1 The Data

3.1.1 Source of the data

The data on which this thesis research is based on is the 2004 Behavioral Surveillance Survey (BSS) conducted by the Central Statistics Agency (CSA),

According to the CSA survey design report, this survey was designed to collect information for the purpose of identifying risk behaviors among sub-populations in order to track behavior change, identify priorities for planning HIV prevention and control programs, and provide key information for advocacy and policy making. It was also aimed to provide information to major program inputs and evidence for the relative success of HIV prevention effects taking place in selected sites.

The survey was conducted in all regions of the country. Both quantitative and qualitative methods were applied in the analysis as it was done in the first round. In this survey extensive mapping procedures were used to identify target groups.

The survey used representative samples of uniformed services (Defense Forces and Police Forces), Road Construction Workers (RCW), Long Distance Truck Drivers (LDTDs), teachers, In-school Youth (ISY), Antenatal care services, Catchments area population, Factory Workers, Pastoralists and cross border community.

The BSS–Ethiopia used sampling strategies that are crucial to the measurement of trends over time. Some target groups were not easily accessible through conventional household or institutional sampling techniques. For most target groups, two-stage sample designs were used. Primary units (kebele, villages, classes, taxi routes) chosen at the first stage and individual respondents were chosen at the second stage. When members of the target group were associated with sites in a fixed manner, conventional clusters were used.

When the target groups were ‘floating’, time-location clusters and targeted snowball sampling (This sampling technique is often used in hidden populations which are difficult for researchers to access; example populations would be drug users or commercial sex workers) were used.

The probability-sampling method was used for the selection of respondents. The sampling frames for the household surveys and institution-based surveys were simple and readily available. However, sampling frame development required preliminary qualitative research and some level of mapping (social and geographic mapping). The objectives of the BSS mapping were to define target groups and identify their locations, to estimate sizes of the potential target groups and systematically list ‘clusters’ of groups to serve as a sampling frame for the main survey. The mapping also aimed to provide information on potential target groups for immediate commencement of intervention activities by governmental and non-governmental agencies, which were planning to operate in the area. The BSS team used various methods for mapping, including interview of key informants, observation by walking through the community, review of records and reports, and discussion with experts and program managers involved in STI/HIV/AIDS control.

This thesis uses data from this survey for analysis of the demographic and socioeconomic factors that influence condom use in four regions: Addis Ababa, SNNPR, Amhara, and Oromia.

3.2 Variables in the Study

Variables to be included in this study are selected from the “2004 Behavioral Surveillance Survey for HIV/AIDS/STIs, Round Two” conducted by the Central Statistics Agency of the Government of Ethiopia.

3.2.1 Dependent Variable

The response/dependent variable in this study, which is ‘ever-used condom’, is dichotomized as 0 if the respondent used condom before the date of the data collection and as 1 if the respondent did not use condom before the date of data collection.

3.2.2 Independent Variables

Predicting whether an event will or will not occur and identifying the variables in making the prediction is an important step in carrying out the study. The independent variables/factors that are used in the study were classified as demographic, socio-economic covariates and other variables. Variables such as age, sex, etc...are considered as demographic variables, while others like economic status of the respondent are considered as socio-economic variables.

Table 3.1 Independent variables with their label and category

Variables	Variable Label	Category
Region	Region where the respondent lives in	3=Amhara 4=Oromia 7=SNNPR 14=A.A
Sex	Sex of the respondent	0=Male 1=Female
Agegroup	Age group of the respondent	0=15-29 1=30-49 2=50 and above
MarStat	Marital Status	0=Single 1=Married 2=Divorced 3=Widowed
Educs	Educational level completed	0=illiterate 1=Read and write 2=Grade 1-12 3=Above grade 12
Relig	Religious affiliation	0=Christian 1=Muslim

		2=Traditional beliefs
MaOccu	Major occupation of the respondent	1=Legislator 2=Professional 3=Unemployed
Income	Income from main work of the respondent	0=<600 1=601-2000 2=Above 2000 3= Unemployed
No.sexInter	Number of sexual intercourse since the beginning of data collection	0=1-5 1=6-20 2=Above 20
FriAlco	Frequency of alcohol drinking	0=Every day 1=Sometimes 2=never
ConNotConf	Condom not comfortable initiating	0=No 1=Yes
UseOthCont	Used other contraceptive	1=Yes 0=No
DuLaKnAppCon	Lake of Knowledge of Appling Condom?	1=Yes 0=No
AwePIPeSoCon	Awareness of place or person of source of condom	1=Yes 0=No
EvHeDiSTI	Ever-heard of diseases that can be transmitted through sexual intercourse	1=Yes 0=No
ComMessHIVAIDSRadTelv	Comment on the messages on HIV or AIDS delivered by radio/television	0=Very clear 1=Clear 2=Somewhat clear 3=Not clear

3.2.3 Interaction

In this study possible interactions Educ*MaOccu, Educ*Income, Income* FriAlco, that is the interaction between educational level and major occupation, educational level and income, and frequency of alcohol drinking and income were considered.

3.3 Methodology

There are different types of multivariate statistical techniques that can be used to predict a binary dependent variable from a set of independent variables. Among these, multiple linear regression analysis and discriminant analysis are two related techniques that quickly come to mind. However, these techniques bring difficulties when the dependent variable is dichotomous or categorical.

It is unreasonable to assume that error is normal and it is difficult to interpret predicted values as probabilities.

Even if linear discriminant analysis does allow direct prediction of group membership, the assumption of multivariate normality of the independent variable as well as equal variance-covariance matrices in the two groups is required for the prediction rule to be optimal.

Among these techniques the logistic regression model is used for the data analysis in this study.

3.4 Logistic Regression

For a binary response Y_j and quantitative explanatory variable X_{ij} , $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$ let $\pi_j = P(X_{ij})$ denote the “success probability” when X_{ij} takes the values x_{ij} . The problem with linear model is that the probability model $E(y) = X\beta$ (where β is the vector of parameters to be estimated) is used to approximate a probability value, $\pi_j = P(Y_j=1)$ within the interval 0 and 1, while $E(Y_j=1)$ is not so constrained. Therefore, we apply the

logit transformation where the transformed quantity $\ln(\pi_j/(1-\pi_j))$ lies in the interval $(-\infty, \infty)$ and it is modeled as

$$\log it(\Pi_j) = \ln\left(\frac{\Pi_j}{1-\Pi_j}\right) = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \dots + \beta_m X_{mj}$$

This implies

$$\Pi_j = \frac{\exp(\beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \dots + \beta_m X_{mj})}{1 + \exp(\beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \dots + \beta_m X_{mj})}$$

where the parameter β_i determines the rate of increase or decrease of X_{ij} on the log of odds that $Y_j=1$, controlling for other X 's. Furthermore, $\exp(\beta_i)$ is the multiplicative effect on the odds of unit increase in X_{ij} , at fixed levels of the others X 's.

In the linear regression we estimate the parameters of the model using the method of least squares. That is, we select regression coefficients that result in the smallest sums of squared distances between the observed and the predicted values of the dependent variables. In logistic regression the parameters of the model are estimated using maximum-likelihood method. That is, the coefficients that make our observed results most'' likely'' are selected. Since the logistic regression model is non-linear, an iterative algorithm is necessary for parameter estimation.

3.5 Parameter Estimation and Statistical Inference

Maximum likelihood estimation is used for estimating the parameters in the logistic regression model. The likelihood L (developed below) is maximized in order to achieve better estimates. The higher the maximized value of L , the better the fit of the model. This is assessed on a log scale by computing $-2\log L$, called-2LL. (This criterion corresponds to residual mean square, *i.e.*, sum of squared errors, in normal multiple

regression models.) When there are several explanatory variables, different models can be assessed using -2LL as a figure-of-merit.

Here is an explanation of maximum likelihood estimation in logistic regression models. We begin with the simpler case of a sample of Bernoulli variables. If we have a sample $y_1, y_2, \dots, y_n$

of 0,1 variables with success probability π , then the log likelihood can be written

$$L = y_1 \ln \pi + (1-y_1) \ln(1-\pi) + y_2 \ln \pi + (1-y_2) \ln(1-\pi) + \dots + y_n \ln \pi + (1-y_n) \ln(1-\pi).$$

The maximum likelihood estimator is the value π of which maximizes this; it turns out to be simply the sample proportion of 1's,

$$(y_1 + y_2 + \dots + y_n)/n.$$

In the logistic regression model,

$$\pi_x = 1/[1 + \exp(-\beta_0 - \beta^T \mathbf{x})],$$

where β_0 is the constant and β is the vector of logistic regression parameters, to be estimated. This is done by maximizing the likelihood by numerical methods.

Statistical inference of the model parameter helps us to judge the significance and magnitude of the effects. Stating the null hypothesis one can test the significance of the effect of X on the binary response

$H_0: \beta = 0$ (the probability of the presence is independent of X)

$H_1: \beta \neq 0$ (the probability of the presence is not independent of X)

For large sample size, assuming normality of maximum likelihood estimates (i.e., the Wald statistic $(Z^2) = [\hat{\beta} / se(\hat{\beta})]^2$ has a chi-squared distribution with one degree of

freedom) it is possible to test the hypothesis under consideration. But the Wald statistic, takes the right tail chi-squared probability above the observed value as its P-values. On the other hand, the likelihood ratio test uses the likelihood function through the ratio of the maximum of the function over the possible parameter values given by the null hypothesis ($\beta = 0$) to the maximum over the large set of possible parameter values for the full model, permitting the null or the alternative hypothesis to be true.

However, for large sample size both the Wald and likelihood ratio test statistics have the same behavior (Agresti, 1996). The Wald statistics has the undesirable property that when the absolute value of the regression coefficients becomes large, the estimated standard error becomes too large. This produces a value of the Wald statistic that is too small, leading to fail to reject the null hypothesis that the coefficient is 0, when in fact we should. Therefore, whenever we have a large coefficient, we should not depend on the Wald statistic for hypothesis testing. Instead we should build a model with and without that variable and base our hypothesis test on the difference between the two likelihood-ratios chi-squared. Hence this study uses the ML ratio test statistics. The ML ratio test statistic is more powerful and reliable for sample of sizes used in practice (Agresti, 1996).

3.6 Model Selection

In logistic regression, as in other multivariate statistical techniques, one may want to identify subset of independent variables that are good predictors of the dependent variable. All the problems associated with variables selection algorithms in regression and discriminant analysis are found in logistic regression as well.

Several model selection procedures exist, none of which is the "best". The cautions that apply to ordinary regression modeling of normal data hold for any generalized linear regression model. For instance, a model with several predictors has the potential for multicollinearity. Strong correlation's among predictors makes it seem that no one variable is important when all the others are in the model. A variable may seem to have a little effect simply because it "overlaps" considerably with other predictors in the model.

The logistic regression procedures have several methods available for model selection. Variables can be entered into the model using Enter, Fstep, Bstep and etc. A better, but computationally more intensive criterion for determining variables to be removed from the model is the likelihood-ratio (LR) test. If the null hypothesis is true and the sample size is sufficiently large, the quantity $-2\text{Log}(L_0/L_1) = -2[\text{Log}(L_0) - \text{Log}(L_1)]$ where L_0 is maximized value of the likelihood function for the simpler model and L_1 is the maximized value of the likelihood function for the full model has a chi-square distribution with k degrees of freedom, where k is the difference between the number of terms in the full model and reduced model. (The model chi-square and the improvement chi-square are both likelihood-ratio tests). When the likelihood-ratio test is used for removing terms from a model, its significance level is compared to the cut off value.

In our case for selection of the model the forward stepwise is used to select those variables which best explain the data. It starts with a model that contains only the constant. At each step, the variable with the smallest significance level for the score statistic, provided it is less than the chosen value, 0.05 is entered into the model.

The score statistic, in the univariate case, is

$$ST = \frac{\sum_i x_i (y_i - \bar{y})}{\sqrt{\bar{y}(1 - \bar{y}) \sum_i (x_i - \bar{x})^2}}$$

Comparison can be done between different models fit to the data in terms of their predictive power. In general linear models the correlation R between the observed responses (Y_j) and the model's fitted value describes this, where R represents the multiple correlations between the response variable and the predictors; the R^2 describes the proportion of the variation in Y that is explained by the predictors. In the case of logistic regression, R is the correlation between binary y observations on the response and the predicted probability π_i . For such models, R is the crude index of predictive power, and it

does not have the nice properties it has for normal general linear model. Nevertheless, R is useful for comparing fits of different models to the same data set (Agresti, 1996)

3.7 Model Checking

After describing and inference about the effects of predictors on a binary response in logistic regression, it is a standard procedure to check that a particular model is appropriate for the data or not.

Fitting logistic regression models provides predicted probabilities that $Y=1$. At each setting of the explanatory variables, one can multiply the predicted probability by the number of subject to obtain a fitted count. The test of the null hypothesis that the model holds compares the fitted and observed counts using a Pearson χ^2 or likelihood-ratio test statistics (G^2). For a fixed number of settings, when most fitted counts are five and more, χ^2 and G^2 have approximately chi-squared distributions. The degree of freedom (the residual df) for the model, equal the number of sample logit (i.e. the number of settings of explanatory variables) minus the number of model parameters. The large χ^2 or G^2 values indicate lack of fit of the model, and P-value is the right-tail probability above the observed value.

When the fit is poor, residuals and other diagnostic measures describe the influence of individual observations on the model fit and highlight reason for the inadequacy (Agresti, 1996)

$$\chi^2 = \sum \frac{(Observed - Fitted)^2}{Fitted}$$

$$G^2 = 2 \sum \left[(Observed) \log \left(\frac{Observed}{Fitted} \right) \right] = -2 \ln \left\{ \frac{LLC}{LLS} \right\}$$

In the second equation

LLC=Maximum likelihood of constant model, and

LLS=Maximum likelihood of full model.

Testing a reduced model against a full model is based on $-2LL$, where LL is the natural log of the maximized likelihood. It is based on the fact that $(-2LL)_{full} - (-2LL)_{reduced}$ is for large distributed approximately as chi-square with the number of d.f. equal to $f_{full} - r_{reduced}$, where these k 's are the numbers of explanatory variables in the two models (see section 3.5).

A constant model is the model which contains only the constant term, while a full is a model which contains all factors selected by the stepwise procedure. Goodness of fit statistics such as G^2 and X^2 are summery indicators of the overall quality of fit. Additional diagnosis analysis is necessary to describe the nature of any lack of fit.

It is important to examine the adequacy of the resulting model in logistic regression. There are comparable diagnostics that should be used to identify data problems. The logistic regression provides a variety of such statistics.

The residual is the difference between the observed probability and the predicted probability of the event based on the model. It can be viewed as a component of $-2LL$, which compares a model to the “perfect” model.

For each case the deviance compares the predicted probability of being in the correct group based on the model to the prediction of $Y=1$.

$$Deviance = \sqrt{-2 \log \left(\frac{L_o}{L_1} \right)}$$

where L_1 is always 1, since the likelihood of the correct prediction in a perfect model is 1, and L_0 in the predicted probability of membership in the correct group. Large values for deviance indicate that the model does not fit the case well. For large sample sizes the deviance is approximately normally distributed. The studentized residual for a case is the change in the model deviance if the case is excluded.

Leverage values are used for detecting observations that have a large impact on the predicted values. Unlike linear regression, the leverage values in logistic regression depend on the dependent variable scores and design matrix and bounded by 0 and 1.

The test statistic is the leverage of the i^{th} case, h_i diagonal element of the matrix

$$\hat{V}^{1/2} X(X' C \hat{V} X)^{-1} X' \hat{V}^{1/2}$$

where

$$\hat{V} = \text{Diag}\{\hat{\pi}_1(1-\hat{\pi}_1), \dots, \hat{\pi}_n(1-\hat{\pi}_n)\}$$

$X = n \times p$ matrix with elements of X_{ij} , the observed value of the i^{th} case of the j^{th} parameter.

$C =$ the asymptotic covariance matrix for $\hat{\beta}$.

Their average value is k/n where k is the number of estimated parameters, including the constant, and n is the sample size.

Cook's distance is a measure of the influence of a case. It tells us how much detecting a case affects not only the residual for that case. Cook's distance depends on the standardized residual for a case, as well as its leverage.

$$\text{Cook's } Di = \frac{Z_i h_i}{(1 - h_i)^2}$$

where z_i is the standardized residual and h_i is the leverage.

The Hosmer-Lemshow test is used to assess goodness of fit of the model. In this approach, data are divided into 10 groups, and from each of such groups the observed and the expected number of events will be computed.

The test statistic is given by:

$$\bar{C} = \sum_{k=1}^{10} \frac{(O_k - E_k)^2}{V_k}$$

where O_k =observed number of events in k^{th} group, E_k =expected number of events in k^{th} group and V_k =a variance correction factor for the k^{th} group.

In summary the model diagnostics indicated above like the residuals (used to detect other inadequacies), the leverage (used for detecting observation that have a large impact on the predicted values), and Cook's distance (a measure of the change in the logits when the case is detected from the model) are useful for model diagnostics (Montgomery and Peck, 1992).

CHAPTER IV

STATISTICAL DATA ANALYSIS

4.1 Introduction

The purpose of this chapter is to measure the effect of the different demographic and socio-economic determinants of condom use in Addis Ababa, Oromia, Amhara, and SNNPR. An attempt has been made to determine the factors that influence the tendency of individuals towards condom use. The research is made using the 2004 Behavioral Surveillance Survey (BSS) data. The BSS used sampling strategies that are essential to the measurement of trends over time. Some target groups were not easily accessible through conventional household or institutional sampling techniques. For most target groups, two-stage sample designs were used. Primary units (clusters) were chosen at the first stage and individual respondents were chosen at the second stage. When members of the target group were associated with sites in a fixed manner, conventional clusters were used. When the target groups were 'floating', time-location clusters and targeted snowball sampling were used. The probability-sampling method was used for the selection of respondents. The sampling frames for the household surveys and institution-based surveys were simple and readily available. After removal of observations with missing values, a total of 2713 observations were included in the analysis. As it was mentioned in Section 1.3, the individuals included in the sample were only legislators, professionals, and unemployed people in the age group 15-59 years. And the results we discuss in this study will focus on these target groups.

4.2 Variables Selection

4.2.1 Bivariate Findings

To determine the factors which are significantly correlated with the dependent variable, a preliminary assessment was used using the chi-square test. The results in Table 4.1 show that the variables selected for the study were strongly associated with the dependent

variable Ever-used condom. From the table all explanatory variables are significant except the variable “condom not comfortable for sex”.

Table 4.1 Bivariate association between Ever-used condom and independent variables

Independent variables	Description	d.f	Chi- Square Value	Sig. (p-value)
Region	Region where the respondent lived	3	118.497	0.000
Sex	Sex of the respondent	1	23.265	0.000
Agegroup	Age group of the respondent	2	47.830	0.000
MarStat	Marital Status	3	167.349	0.000
Edus	Educational level completed	3	209.979	0.000
Relig	Religious affiliation	2	26.587	0.000
MaOccu	Major occupation of the respondent	2	16.35	0.000
Income	Income from main work of the respondent	3	39.827	0.000
ConNotConf	Condom not comfortable initiating	1	1.677	0.195
No.sexInter	Number of sexual intercours	3	61.574	0.000
FriAlco	Frequency of alcohol drinking	2	9.281	0.010
UseOthCont	Used other contraceptives	1	4.832	0.028
DuLaKnAppCon	Due to lack of knowledge of applying condom	1	13.362	0.000
AwePIPeSoCon	Awareness of place or person of source of condom	1	104.148	0.000
EvHeDiSTI	Ever heard of diseases that can be transmitted through sexual intercourse	1	25.616	0.000
ComMessHIVAI DSRadTelv	Comment on the messages on HIV or AIDS delivered by	3	18.441	0.000

	radio/television			
--	------------------	--	--	--

4.2.2 Multivariate Findings

The main problem with the bi-variate approach is that it ignores the possibility that a collection of variables, each of which could be weakly associated with the outcome, can become an important predictor of the outcome when taken together (Hosmer and Lemeshow, 1989). Hence, multivariate logistic regression approach that takes into account the drawback mentioned by the bi-variant technique is considered in the following analysis.

Using this method, the model that best describes the dependent variable Ever-used condom is fitted using the explanatory variables. The forward stepwise technique is used to select the best model. The result of logistic regression model is given in Table 4.3. The final (optimal) logistic regression model includes only those significant variables.

Table 4.2 Categorical Variables Coding

		Frequency	Parameter coding		
			(1)	(2)	(3)
Number of Sexual Intercourse since begin of data collection	1-5	1254	1.000	.000	.000
	6-20	398	.000	1.000	.000
	21-30	389	.000	.000	1.000
	>30	672	.000	.000	.000
Marital Status	Single	436	1.000	.000	.000
	Married	2172	.000	1.000	.000
	Divorced	57	.000	.000	1.000
	Widowed	48	.000	.000	.000
Educational Level Completed	Illiterates	1220	1.000	.000	.000
	Read and write	498	.000	1.000	.000
	Grade 1-12	820	.000	.000	1.000

	Above grade 12	175	.000	.000	.000
Income from Main Work of the Respondent	<600	1332	1.000	.000	.000
	601-2000	108	.000	1.000	.000
	>2001	5	.000	.000	1.000
	Unemployed	1268	.000	.000	.000
Region in which the Respondent lives	Amhara	736	1.000	.000	.000
	Oromia	1252	.000	1.000	.000
	SNNPR	359	.000	.000	1.000
	A.A	366	.000	.000	.000
Comment on the Messages on HIV or AIDS Delivered by Radio/Television	Very clear	707	1.000	.000	.000
	Clear	1864	.000	1.000	.000
	Somewhat clear	113	.000	.000	1.000
	Not clear	29	.000	.000	.000
Major Occupation of the Respondent	Legislator	18	1.000	.000	-
	Professional	1467	.000	1.000	-
	Unemployed	1228	.000	.000	-
Age Group	15-29	1198	1.000	.000	-
	30-49	1320	.000	1.000	-
	50 and older	195	.000	.000	-
Religious affiliation	Christian	1695	1.000	.000	-
	Muslim	960	.000	1.000	-
	Traditional	58	.000	.000	-
Frequency of alcohol drinking	Every day	25	1.000	.000	-
	some times	852	.000	1.000	-
	Never	1836	.000	.000	-
Ever-heard diseases that can be transmitted through sexual	No	343	1.000	-	-
	Yes	2370	.000	-	-

intercourse					
Sex of the Respondent	Male	1426	1.000	-	-
	Female	1287	.000	-	-
Used other Contraceptive	No	2620	1.000	-	-
	Yes	93	.000	-	-
Condom not comfortable initiating	No	2704	1.000	-	-
	Yes	9	.000	-	-
Awareness of place or Person that is the source of condom	No	895	1.000	-	-
	Yes	1818	.000	-	-
Lake of Knowledge of Appling Condom?	No	2521	1.000	-	-
	Yes	192	.000	-	-

Based on these coding schemes given in Table 4.2 the logistic regression coefficients can be estimated using the maximum likelihood estimation method. This will be done using the SPSS package.

Table 4.3 Estimates for the final Logistic Regression Model

	B	S.E.	Wald	df	Sig.	Exp(B)
Region			22.981	3	.000	
Region(1)	.257	.188	1.862	1	.172	1.292
Region(2)	.934	.205	20.771	1	.000	2.546
Region(3)	.603	.344	3.077	1	.079	1.828
sex(1)	-.801	.159	25.553	1	.000	.449
agegroup			31.104	2	.000	
agegroup(1)	-1.407	.422	11.136	1	.001	.245
agegroup(2)	-.606	.420	2.084	1	.149	.545

MarStat			35.410	3	.000	
MarStat(1)	-.184	.370	.249	1	.618	.832
MarStat(2)	.815	.366	4.967	1	.026	2.260
MarStat(3)	.520	.538	.934	1	.334	1.682
Edus			44.040	3	.000	
Edus(1)	1.145	.285	16.179	1	.000	3.144
Edus(2)	1.354	.297	20.769	1	.000	3.875
Edus(3)	.101	.219	.211	1	.646	1.106
UseOthCon(1)	.786	.317	6.159	1	.013	2.195
AwePIPeSoCon(1)	1.437	.321	20.016	1	.000	4.207
Constant	1.138	.667	2.908	1	.088	3.121

- a Variable(s) entered on step 1: Edus.
- b Variable(s) entered on step 2: MarStat.
- c Variable(s) entered on step 3: AwePIPeSoCon.
- d Variable(s) entered on step 4: agegroup.
- e Variable(s) entered on step 5: sex.
- f Variable(s) entered on step 6: Region.
- g Variable(s) entered on step 7: UseOthCon.

Variables in the equation are entered Educational level completed, Marital Status, Awareness of place or person which is the source of condom, Age group, Sex, Region, Use other contraceptive method in step one, step two, step three, step four, step five, step six, step seven, respectively. The interaction terms are insignificant to the model as confirmed by the results in the Annex.

The significance of the Wald statistics (under the column with heading Sig.) indicates the importance of the predictor variables in the model. A high value of the Wald statistics shows that the corresponding predictor variable is significant.

4.3 Model checking and diagnostics

A model should be assessed and diagnosed for model adequacies. The chi-square goodness of fit statistic and the classification power for the criterion variables (i.e. Ever-used condom) are given in Table 4.3. A good model is one that results in the high likelihood of the observed results. This translates to small values of -2LL. If a model fits perfectly, the likelihood is 1 (meaning -2LL is 0). Table 4.3 shows, based on the wald statistics, almost all of the coefficients (β 's) significant at a one percent level of significance.

Goodness of fit of the model can be

(i) The classification table shows that of the 2713 respondents included in the analysis, 91.1% were correctly classified. This clearly shows that the final logistic model performs well (Table E).

(ii) -2LL statistic for the logistic model that contain only the constant and the final logistic regression model.

The -2log likelihood (-2LL) statistic for the logistic model that contains only the constant, -2LL is 1709.220 (see Table 4.4).

Table 4.4 -2LL for the model contain only the constant

-2Log likelihood	Coefficients Constant
1709.220	2.249

a. Constant is included in the model.

b. Initial -2 Log Likelihood: 1709.220

The goodness of fit statistics for the final logistic regression model (see Table 4.5) gives for the final model a -2Log likelihood value 1321.885.

Table 4.5 -2LL for the final logistic regression model

-2 Log likelihood	df	Sig.
1321.885	14	.000

There is a significance difference in -2LL between the final model and the model with only the constant (the model chi-square) at $\alpha = 0.05$ level. This indicates that the variables in the final model explain the data quite well.

iii) The Hosmer-Lemeshow test is also another way of assessing goodness of fit of the model

Table 4.6 Hosmer-Lemeshow Test

Chi-square	df	Sig.
10.792	8	.214

As we can see from Table 4.6, we do not reject H_0 at $\alpha=0.05$ level of significance. This shows that there is no sufficient evidence to reject the null hypothesis and it confirms that our model has a good fit.

By using the classification table, chi-square goodness of fit test and Hosmer-Lemeshow test we can say the fitted model is statistically satisfactory.

The residual plots, that is, the graph of the residuals versus the explanatory variables showed the presence of some outliers which will influence the estimates. Thus, the outliers were removed from the analysis.

Diagnosis

(i) Standardized residual

The standardized residuals for the final regression model are small. Those values 88.6 percent of them lies between -1 and 1 while 97.4 percent lie between -2 and 2 (See Table 4.7).

Table 4.7 Standard Residual

		Frequency	Percent	Valid Percent	Cumulative Percent
Standard Residual	Less than -2	71	2.6	2.6	2.6
	-2 to -1	181	6.7	6.7	9.3
	-1 to 1	2405	88.6	88.6	97.9
	1 to 2	56	2.1	2.1	100.0
	Greater than 2	0	0.0	0.0	100.0
	Total	2713	100.0	100.0	-

(ii) The deviance

For the final logistic regression model, we also observed small values for the deviance. Those values that lie between -1 and 1 accounted for 89.4 percent while 97.4 lie between -2 and 2. (See Table 4.8).

Table 4.8 Deviance

		Frequency	Percent	Valid Percent	Cumulative Percent
<i>Deviance</i>	Less than -2	71	2.6	2.6	2.6
	-2 to -1	181	6.7	6.7	9.3
	-1 to 1	2426	89.4	89.4	98.7
	1 to 2	35	1.3	1.3	100.0
	Greater than 2	0	0.0	0.0	100.0
	Total	2713	100.0	100.0	-

(iii) Cook's distance

For the final logistic regression model, we also observed small values for the Cook's distance. Those values and less than one accounted for 100 percent. (See Table 4.9).

Table 4.9 Cook's Distance

		Frequency	Percent	Valid Percent	Cumulative Percent
Cook's Distance	Less than 1	2713	100.0	100.0	100.0
	Greater than 1	0	0.0	0.0	100.0

iv) Leverage Value

Leverage values are used for detecting observations that have large impact on the predicted values. Those values greater than 0 and less than one accounted for 100 percent. (See Table 4.10).

Table 4.10 Leverage Value

		Frequency	Percent	Valid Percent	Cumulative Percent
Leverage Value	Less than 0	0	0.0	0.0	0.0
	0 to 1	2713	100.0	100.0	100.0
	Greater than 1	0	0	0	100.0

Small values for standardized residual, deviance, leverage value and Cook's distance indicate that the final logistic regression model fit the data quite well.

CHAPTER V

DISCUSSION, CONCLUSION AND RECOMMENDATION

5.1. Discussion

The result of Bivariate finding shows region where the respondent lived, sex of the respondent, age group of the respondent, marital Status, educational level completed, religion of the respondent, major occupation of the respondent, income from main work of the respondent, number of sexual intercours, frequency of alcohol drinking, lack of knowledge of applying condom, awareness of about source of condom, ever heard of diseases that can be transmitted through sexual intercourse, comment on the messages on HIV or AIDS delivered by radio/television show a significant association with condom use.

The result of logistic regression analysis show the determinants of condom use (Ever-used condom) as given in Table 4.3. The results of the logistic regression analysis indicate that the determinants of not using/using condom during sexual intercourse include educational level completed, region, sex, age, marital status, use of other contraceptive method, and awareness of place or person that is the source of condom.

When we see each factor closely, we find that people living out of Addis Ababa have fewer tendencies to use condom. Compared to A.A, people living in Oromia and SNNPR regions are about three and two times, more likely not to use condom respectively. Amhara region shows some kind of similarity with A.A towards condom use (odds ratio=1.29).The female gender also has shown lower tendency towards condom use. People at older ages are also less likely to use condom. This indicated by the odds of not using condom for the age group 15-29 and 30-50 which is 0.24 and 0.55 respectively. Singles appear to be 83.2% less likely with those windowed. Married and divorced individuals are about two times more likely not to use condom when compared to the windowed group.

People who are illiterate and who can only read and write are about three and four times more likely not to use condom respectively when compared with those who are above grade 12. And those whose educational status lies between grade 1 and grade 12 have similar behaviour towards condom use with those people who are more than grade 12 (odds ratio=1.11). Those who don't use other contraceptives are twice as likely not to use condoms as compared to those who use other contraceptive. Finally we find that those who don't know where they can find condom are about 4 times more likely not use condom compared to those who have the idea.

The interactions of educational level and major occupation, educational level and income, and frequency of alcohol drinking and income are not associated with the dependent variable as confirmed by the results in the Annex.

5.2. Conclusion

The logistic regression results indicated some determinants of condom use. Educational level, marital status, awareness of the person or people about sources of condom, age group, sex, region, use other contraceptive method are identified as the significant factors/variables. The results of the analysis imply that strategies and policies aimed at raising awareness about condom use, higher educational status, increasing the source of condom, and also the awareness of contraceptive method must be developed.

5. 3. Recommendation

Condom is the basic means for protecting HIV/AIDS and other sexually transmitted diseases (STDS).Condom is not only protecting against STDS but also against unwanted pregnancy. The need for condom is growing as HIV/AIDS and other STDS spread. We will recommend for increasing condom use, strategies and policies focused on:

- Educating people in order to create awareness about condom use.
- Encourage hotels, shops, supermarkets, etc to distribute condoms.

- Encourage governmental/non governmental organization to participate in the protection of HIV/AIDS and other STDS by using condom.

Reference

- Agresti, A.** (1996), *An Introduction to Categorical Data Analysis*. New York, Wiley.
- Arjun, C. and Ipshita, R.** (2006), The effect of AIDS awareness on condom use intention among truck drivers in India. *Journal of Marketing Communication*, 10(1):17-33.
- Belachew Gebre** (2002) Factors affecting accessibility and utilization of condom in Ethiopia. *National library of Medicine*, 14(3):24-30.
- Brandy, A. and Laurie, M.** (2002), Review of Research on Adolescent Religiosity and Sexual Behavior, ETR's Resource center for adolescence Pregnancy Prevention, 2(1):65-70.
- Family Health International** (2004).HIV Counseling and Testing for youth. A manual for providers, FHI 2004, Youth Net.
- Family Health International AIDSCAP Program in Ethiopia (FHI/AIDSCAP)** (2000) Condom Promotion and Distribution. Final Report for the Executive Summary, 1(1):13-20.
- Hosmer, D.W. and Lemeshow, S.** (1989): *Applied Logistic Regression*, New York: John Wiley.
- Juarez, F. and Martín, T.C.** (2006) Safe Sex Versus Safe Love? Relationship Context and Condom Use among Male Adolescents in the Favelas of Recife, Brazil. *Archives of sexual behavior*, 35(1):74-85.
- Jegu Degu Alene** Knowledge and practice of condom in preventing HIV/AIDS infection among commercial sex workers in three small towns of northwestern Ethiopia. *Ethiopian Journal of Health development* 2002 16(3): 277-286.
- Manchikanti, A.** (2005) The Relationship between Sexual Violence and Reproductive Health among Female Youth in Colombia. *International Family Planning Perspective*, 30 (4):165-173.
- Mannin,W.D Longmore,M.A and Giordano,P.C** (2000) The Relationship Context of Contraceptive Use at First Intercourse. *Family Planning Perspectives*, 32(3):125-140.

- Mesfin Samuel, Rahel Adamu, and Sara Ingidashet** (2004) Psychosocial and Contextual Determinants of Past and Intended Condom Use among Ethiopian Secondary School Students. *Northeast African Studies*, 7(1): 85-108.
- Ministry of Health** (2004). AIDS in Ethiopia; Disease Prevention and Control Department. 5th edition: Addis Ababa, Ethiopia.
- Ministry of Health** (2002). HIV/AIDS Behavioral Surveillance Survey (BSS), Ethiopia. Round one.
- Ministry of Health** (2005). AIDS in Ethiopia. Disease Prevention and Control Department. 6th edition: Addis Ababa, Ethiopia.
- Ministry of Health** (2004). HIV/AIDS/STIs Behavioral Surveillance Survey, Round Two (BSS): Addis Ababa, Ethiopia.
- Mnyika, K.S. Klepp, K.I Kvale, G. and Ole-King'Ori, N.** (1997) Determinants of high-risk sexual behavior and condom use among adults in the Arusha region, Tanzania. *International Journal of STD & AIDS*, 8(3): 176-183.
- Montgomery, D.C., and Peck, E.A. (1992).** Introduction to Linear Regression Analysis, Wiley, New York.
- Muhwava, W.** (2004) Condom Use within Marriage and Consensual Unions in the Era of HIV/AIDS in Zimbabwe. *African Population Studies* 19(1):122-125.
- Mumtaz, Z. Slaymaker, E. and Salway, S.** (2004) condom use in uganda and zimbabwe: exploring the influence of gendered access to resources and couple-level dynamics. A focus on gender: Collected papers on gender using DHS data. Calverton, Maryland, USA. 12 (3): 119-244.
- Pullum, T. Cleland, J. and Shah, I.** (2004) Factors Affecting Condom Use in Primary Sexual Relationships in Southern and Eastern Africa. Prepared for 2004 Annual Meeting of the Population Association of America, Boston.
- Taffa, N. Klepp, I. Sundby, T. and Bjune, G.** (2002) Psychosocial determinants of sexual activity and condom use intention among youth in Addis Ababa. *International Journal of STD & AIDS*, 13(10): 714-719.
- United Nations HIV/AIDS (UNAIDS).** Report on the Global AIDS Epidemic, 4th Global Reports, Geneva, Switzerland, 2004.

United Nations AIDS/HIV (UNAIDS). AIDS Epidemic update, Special reports on HIV Prevention. Dec. 2005.

Annex

Table A Correlation Matrix

		Constant	Edus(1))	Constant	MarStat(1)	Edus(1)	Constant	MarStat(1)	Edus(1)	AwePIPeSoCon(1)	agegroup(1)	sex(1)	Region(1)	UseOtherCon(1)
Step 1	Constant	1.000	-.719											
	Edus(1)	-.719	1.000											
	Edus(2)	-.652	.469											
	Edus(3)	-.899	.646											
Step 2	Constant			1.000	-.837	-.319								
	MarStat(1)			-.837	1.000	-.035								
	MarStat(2)			-.856	.909	-.057								
	MarStat(3)			-.575	.611	-.014								
	Edus(1)			-.319	-.035	1.000								
	Edus(2)			-.294	-.033	.485								
	Edus(3)			-.456	.005	.652								
Step 3	Constant						1.000	-.838	-.298	-.011				
	MarStat(1)						-.838	1.000	-.043	.005				
	MarStat(2)						-.856	.910	-.061	-.010				
	MarStat(3)						-.578	.614	-.019	.011				
	Edus(1)						-.298	-.043	1.000	-.236				
	Edus(2)						-.294	-.033	.481	-.055				
	Edus(3)						-.454	.004	.634	-.009				
	AwePIPeSoCon(1)						-.011	.005	-.236	1.000				
Step 4	Constant						1.000	-.576	-.217	-.009	-.716			
	agegroup(1)						-.716	-.004	.011	.001	1.000			
	agegroup(2)						-.691	-.020	.001	.010	.934			
	MarStat(1)						-.576	1.000	-.040	.005	-.004			
	MarStat(2)						-.618	.908	-.059	-.013	.044			
	MarStat(3)						-.411	.614	-.017	.007	.017			
	Edus(1)						-.217	-.040	1.000	-.235	.011			
	Edus(2)						-.181	-.036	.485	-.057	-.036			
	Edus(3)						-.300	.004	.637	-.012	-.027			
	AwePIPeSoCon(1)						-.009	.005	-.235	1.000	.001			
Step 5	Constant						1.000	-.560	-.257	-.029	-.719	-.311		
	sex(1)						-.311	.031	.149	.051	.162	1.000		
	agegroup(1)						-.719	.003	.042	.013	1.000	.162		

	agegroup(2)						-.684	-.016	.028	.020	.933	.117		
	MarStat(1)						-.560	1.000	-.032	.013	.003	.031		
	MarStat(2)						-.592	.907	-.062	-.007	.046	.010		
	MarStat(3)						-.410	.617	-.009	.013	.027	.047		
	Edus(1)						-.257	-.032	1.000	-.234	.042	.149		
	Edus(2)						-.167	-.043	.485	-.056	-.037	-.011		
	Edus(3)						-.314	.003	.645	-.011	-.009	.085		
	AwePIPeSoCon(1)						-.029	.013	-.234	1.000	.013	.051		
Step 6	Constant						1.000	-.557	-.240	-.024	-.699	-.311	-.181	
	Region(1)						-.181	.065	-.057	-.026	-.044	.021	1.000	
	Region(2)						-.065	.072	-.119	-.036	-.130	-.140	.563	
	Region(3)						-.043	.041	-.190	.009	-.031	-.077	.334	
	sex(1)						-.311	.014	.166	.046	.184	1.000	.021	
	agegroup(1)						-.699	.000	.048	.017	1.000	.184	-.044	
	agegroup(2)						-.676	-.013	.037	.022	.929	.132	.000	
	MarStat(1)						-.557	1.000	-.047	.012	.000	.014	.065	
	MarStat(2)						-.588	.896	-.052	-.004	.068	.026	-.001	
	MarStat(3)						-.410	.612	-.018	.011	.032	.026	.062	
	Edus(1)						-.240	-.047	1.000	-.234	.048	.166	-.057	
	Edus(2)						-.169	-.050	.501	-.060	-.034	.004	.017	
	Edus(3)						-.314	-.002	.641	-.010	-.004	.096	.011	
	AwePIPeSoCon(1)						-.024	.012	-.234	1.000	.017	.046	-.026	
Step 7	Constant						1.000	-.496	-.179	-.006	-.629	-.240	-.150	-.445
	Region(1)						-.150	.065	-.057	-.026	-.039	.009	1.000	-.027
	Region(2)						-.047	.072	-.118	-.037	-.128	-.143	.562	-.023
	Region(3)						-.016	.041	-.190	.011	-.030	-.073	.335	-.052
	sex(1)						-.240	.014	.175	.048	.184	1.000	.009	-.085
	agegroup(1)						-.629	.001	.047	.017	1.000	.184	-.039	.004
	agegroup(2)						-.621	-.013	.034	.021	.929	.129	.003	.034
	MarStat(1)						-.496	1.000	-.046	.013	.001	.014	.065	-.004
	MarStat(2)						-.563	.892	-.057	-.006	.068	.021	-.001	.086
	MarStat(3)						-.382	.610	-.022	.009	.031	.018	.058	.040
	Edus(1)						-.179	-.046	1.000	-.234	.047	.175	-.057	-.078
	Edus(2)						-.133	-.049	.503	-.059	-.035	.009	.019	-.039
	Edus(3)						-.255	-.002	.642	-.009	-.005	.103	.010	-.058
	UseOthCon(1)						-.445	-.004	-.078	-.035	.004	-.085	-.027	1.000
	AwePIPeSoCon(1)						-.006	.013	-.234	1.000	.017	.048	-.026	-.035

Table B Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1503.301(a)	.073	.156
2	1422.955(a)	.100	.214
3	1388.573(b)	.111	.238
4	1370.690(b)	.117	.251
5	1351.662(b)	.123	.264
6	1327.361(b)	.131	.281
7	1321.885(b)	.133	.285

a Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

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

Table C Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	2	1.000
2	1.554	4	.817
3	3.631	5	.604
4	4.118	8	.846
5	7.978	8	.436
6	13.534	8	.095
7	10.792	8	.214

Table D Contingency Table for Hosmer and Lemeshow Test

		Dependent Ever-used Condom = Yes		Dependent Ever-used Condom = No		Total
		Observed	Expected	Observed	Expected	
Step 1	1	40	40.000	135	135.000	175
	2	161	161.000	659	659.000	820
	3	24	24.000	474	474.000	498
	4	34	34.000	1186	1186.000	1220
Step 2	1	103	101.833	174	175.167	277
	2	25	26.241	130	128.759	155
	3	73	72.926	490	490.074	563
	4	19	20.291	213	211.709	232
	5	12	14.900	408	405.100	420

	6	27	22.810	1039	1043.190	1066
Step 3	1	105	101.233	166	169.767	271
	2	33	33.103	165	164.897	198
	3	72	71.666	436	436.334	508
	4	17	24.331	372	364.669	389
	5	20	18.278	453	454.722	473
	6	5	5.428	266	265.572	271
	7	7	4.961	596	598.039	603
Step 4	1	107	102.781	169	173.219	276
	2	52	54.756	238	235.244	290
	3	48	49.181	344	342.819	392
	4	14	16.935	207	204.065	221
	5	12	11.744	216	216.256	228
	6	6	5.536	156	156.464	162
	7	10	8.399	256	257.601	266
	8	2	4.060	201	198.940	203
	9	5	2.851	238	240.149	243
	10	3	2.757	429	429.243	432
Step 5	1	109	102.780	152	158.220	261
	2	43	50.535	208	200.465	251
	3	43	38.239	243	247.761	286
	4	22	25.174	254	250.826	276
	5	12	14.608	236	233.392	248
	6	9	12.569	297	293.431	306
	7	12	7.338	273	277.662	285
	8	5	4.378	323	323.622	328
	9	3	2.413	279	279.587	282
	10	1	.967	189	189.033	190
Step 6	1	119	111.267	158	165.733	277
	2	50	51.138	220	218.862	270
	3	31	35.045	237	232.955	268
	4	18	22.668	249	244.332	267
	5	7	14.777	265	257.223	272
	6	12	9.496	257	259.504	269
	7	8	6.857	263	264.143	271
	8	6	3.971	256	258.029	262
	9	5	2.311	266	268.689	271
	10	3	1.470	283	284.530	286
Step 7	1	117	111.514	154	159.486	271
	2	50	50.961	220	219.039	270
	3	33	35.106	232	229.894	265
	4	17	21.646	240	235.354	257

	5	11	15.271	259	254.729	270
	6	6	8.853	241	238.147	247
	7	9	6.874	246	248.126	255
	8	8	4.467	262	265.533	270
	9	5	2.900	320	322.100	325
	10	3	1.409	280	281.591	283

Table E Classification Table(a)

			Predicted		
			Dependent Ever-used Condom		Percentage Correct
			Yes	No	
Step 1	Dependent Ever-used Condom	Yes	0	259	.0
		No	0	2454	100.0
	Overall Percentage				90.5
Step 2	Dependent Ever-used Condom	Yes	0	259	.0
		No	0	2454	100.0
	Overall Percentage				90.5
Step 3	Dependent Ever-used Condom	Yes	0	259	.0
		No	0	2454	100.0
	Overall Percentage				90.5
Step 4	Dependent Ever-used Condom	Yes	0	259	.0
		No	0	2454	100.0
	Overall Percentage				90.5
Step 5	Dependent Ever-used Condom	Yes	13	246	5.0
		No	11	2443	99.6
	Overall Percentage				90.5
Step 6	Dependent Ever-used Condom	Yes	48	211	18.5
		No	30	2424	98.8
	Overall Percentage				91.1
Step 7	Dependent Ever-used Condom	Yes	48	211	18.5
		No	30	2424	98.8
	Overall Percentage				91.1

a The cut value is .500

Table F Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 7(g)	Region			22.981	3	.000	
	Region(1)	.257	.188	1.862	1	.172	1.292
	Region(2)	.934	.205	20.771	1	.000	2.546
	Region(3)	.603	.344	3.077	1	.079	1.828
	sex(1)	-.801	.159	25.553	1	.000	.449
	agegroup			31.104	2	.000	
	agegroup(1)	-1.407	.422	11.136	1	.001	.245
	agegroup(2)	-.606	.420	2.084	1	.149	.545
	MarStat			35.410	3	.000	
	MarStat(1)	-.184	.370	.249	1	.618	.832
	MarStat(2)	.815	.366	4.967	1	.026	2.260
	MarStat(3)	.520	.538	.934	1	.334	1.682
	Edus			44.040	3	.000	
	Edus(1)	1.145	.285	16.179	1	.000	3.144
	Edus(2)	1.354	.297	20.769	1	.000	3.875
	Edus(3)	.101	.219	.211	1	.646	1.106
	UseOthCon(1)	.786	.317	6.159	1	.013	2.195
	AwePIPeSoCon(1)	1.437	.321	20.016	1	.000	4.207
	Constant	1.138	.667	2.908	1	.088	3.121

a Variable(s) entered on step 1: Edus.

b Variable(s) entered on step 2: MarStat.

c Variable(s) entered on step 3: AwePIPeSoCon.

d Variable(s) entered on step 4: agegroup.

e Variable(s) entered on step 5: sex.

f Variable(s) entered on step 6: Region.

g Variable(s) entered on step 7: UseOthCon.

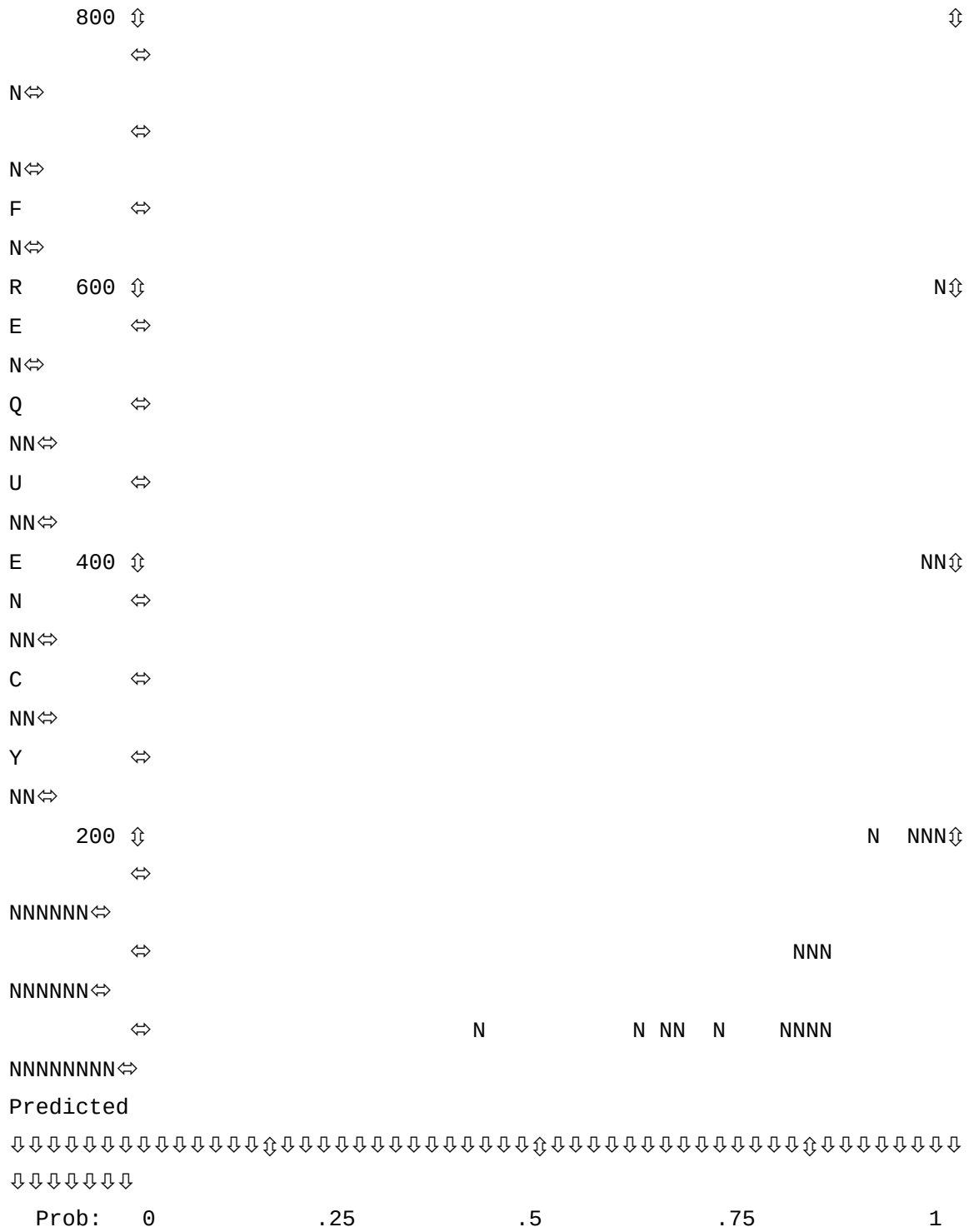
Table G Model if Term Removed

Variable		Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 1	Edus	-854.610	205.920	3	.000
Step 2	MarStat	-751.650	80.346	3	.000
	Edus	-787.321	151.688	3	.000
Step 3	MarStat	-730.671	72.769	3	.000
	Edus	-734.971	81.370	3	.000
	AwePIPeSoCon	-711.477	34.382	1	.000
Step 4	agegroup	-694.286	17.883	2	.000
	MarStat	-714.761	58.832	3	.000
	Edus	-722.910	75.131	3	.000

	AwePIPeSoCon	-702.937	35.185	1	.000
Step 5	sex	-685.345	19.028	1	.000
	agegroup	-687.883	24.104	2	.000
	MarStat	-704.924	58.186	3	.000
	Edus	-710.167	68.672	3	.000
	AwePIPeSoCon	-691.177	30.692	1	.000
Step 6	Region	-675.831	24.301	3	.000
	sex	-676.259	25.156	1	.000
	agegroup	-679.620	31.879	2	.000
	MarStat	-679.211	31.061	3	.000
	Edus	-689.107	50.852	3	.000
	AwePIPeSoCon	-677.475	27.589	1	.000
Step 7	Region	-672.821	23.758	3	.000
	sex	-674.242	26.599	1	.000
	agegroup	-677.398	32.912	2	.000
	MarStat	-678.185	34.485	3	.000
	Edus	-685.496	49.107	3	.000
	UseOthCon	-663.680	5.476	1	.019
	AwePIPeSoCon	-674.160	26.435	1	.000

Step number: 7

Observed Groups and Predicted Probabilities



Group: YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

Predicted Probability is of Membership for No

The Cut Value is .50

Symbols: Y - Yes

N - No

Each Symbol Represents 50 Cases.

DECLARATION

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

Name: Yonas Kassa

Signature:

Place: Faculty of Science, Addis Ababa University

Date: September, 2007

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

.....

Professor Eshetu Wencheko