2/3

School of Graduate Studies
College of Natural Science
Department of Statistics

Determinants of Female Labor Force Participation in Three Major Towns of Amhara Region

By
Fekadu Muluye

A thesis submitted to the School of Graduate Studies of Addis Ababa University in partial fulfillment of the requirements for the Degree of Master of Science in Statistics

March 2011
Addis Ababa

Determinants of Female Labor Force Participation in Three Major Towns of Amhara Region

By

FEKADU MULUYE

Approved by the Board of Examiner:

....

Advisor (prof. Eshetu Wencheko) **Signature**

....

Department Head **Signature**

Butte Gotu ....

Examiner **Signature**

Mekonnen Tadesse ....

Examiner **Signature**

March 2011
Addis Ababa

ABSTRACT

Labor supply, besides being a key component of socio-economic development, is vital for the advancement of women. In this regard, this paper attempts to investigate the determinants of female labor force participation i.e. to identify those factors influencing women's labor supply decision to participate in labor force in three major towns of Amhara region. Analysis of the data from the 2009 Urban Employment and Unemployment Survey suggests that Age, Marital status, and kinship to the household head determine women's labor force participation in the three major towns of the region. Women aged 23-33 years are more likely to participate than other age groups. Women with no education are economically active as those with high school or above high school education in Gondar town. It is also found that being married decreases women LFP in Gondar and Bahardar towns and the presence of children in the household increases women labor force participation in Gondar and Bahardar towns. However, the pattern, the effect, and magnitude of these factors on women's labor force participation depend on the town in which a woman lives.

Table of Contents

reports summary statistics, univariate and multiple findings, and the empirical result interpretations and discussions. Finally Chapter Five presents conclusion and recommendation based on the empirical results.

1.2. Statement of the Problem

The ultimate goal of development efforts of any government is to ensure sustainable economic growth as a means for combating poverty in the country. However, it is necessary to move beyond increasing the national income to ensure that such income is translated into promoting and/or improving the well being of its people by providing or creating employment opportunities in gender equitable manner.

Indeed, the right to employment (whether through wage employment or through self employment) is a basic human right as it provides the means of earning an income that would allow for improving or ensuring a basic standard of living. The challenges facing the Government of Ethiopia in fulfilling the right to employment progressively are twofold: managing the dynamic of population growth and expansion of labor intensive productive activities (United Nation Development Assistance Framework, 2001).

The Government of Ethiopia designed various policies and strategies aimed at increasing employment opportunities. The small scale and micro industry development strategies introduced to support the efforts to expand employment opportunities as key factors for poverty reduction are an examples.

Despite the efforts that have been made so far, employment opportunities are still scarce for both men and women. The 2009 UEUS has indicated that, an increasing trend toward unemployment has been observed in urban areas. In connection with this, women are disadvantaged than men. Women are an important segment of Ethiopian population making up 52% of the total population and 47% the urban labor force with unemployment rate of 30% such that, only 18% of them are employed both in public and private field of the sector against 82% of men.

estimates of this relationship have a negative relationship between the number of children and a woman's labor supply (Brewster and Rindfuss, 2000). The problem with these estimates is that they cannot say anything about causality, because the causal impact of female labor force participation on fertility may occur along a number of complex pathways because both female labor force participation and lower fertility may reinforce each other. The relationships between female labor force participation and fertility have been studied based on the maternal role incompatibility hypothesis, which states that an inverse relationship occurs between women's work and fertility only when the roles of worker and mother conflict (Goldstein, 1972 and Bindary *et al.*, 1973). The implication of this hypothesis is that a negative relationship between female employment and fertility exists to the extent that they are competing uses of time. Otherwise, we should expect to find no relationship, or even a positive relation between employment and fertility. Substantial empirical evidence abounds both in support of and against this hypothesis (Mason and Palan, 1981, Isvan 1991).

The relationship between education and FLFP has been studied extensively by various scholars related to dimensional factors affecting it. In human capital theory, education is the prime determinant of labor market outcomes (i.e. jobs and earnings). Three most popular hypotheses among human capital theories are based on the idea of opportunity cost. Since education is an investment, and since education and earnings potential are positively related, education raises the opportunity cost of not working and thus, increases the incentive to seek employment. For this reason, educational attainment included as a proxy variable for a 'positive test 'or 'pure preference' for market work (Bowen and Finegan, 1969).

However, Standing (1978) argues, the mere fact that education could improve women's employment opportunity does not mean that they will necessarily wish or be able to take advantage of them. He also raises the objection to opportunity cost argument: 1) there may be an inverse relationship between women's education attainment and their LFP, since educated women typically marry educated men who have potentially high earnings (which tend to reduce women's financial incentives to seek employment); 2) within the constraints of the nuclear family, the opportunity cost of non- participation in the labor force may even be lower for an educated woman with children, because an educated woman is probably in a better situations than a less educated women in take care of her child early education (Sheehan and Standing, 1978). They further emphasize that this is particularly likely to be the case in low income countries where institutionalized educational facilities for young children are relatively undeveloped.

Standing (1978) also asserts that education may, to a certain extent, increase the propensity of women to participate in the labor force by raising income aspirations and the degree of dissatisfaction with any given level of family income. However, Standing further contends that since education also raises income and occupational expectations, a positive relationship cannot be presumed because those expectations cannot be always realized. Instead he presents the 'status frustration' effect. The work available to educated women is often of lower status and less well paid than that to which they feel entitled and for which their education qualifies them. In many cases they have to accept jobs inferior to those of men with levels of education comparable to or even inferior to their own. This may lead to a pronounced 'status frustration' effect, whereby women who feel entitled to a certain level of income withdraw from labor force, rather than accept some lower paying, low status job. Standing concludes that this status frustration effect may be particularly common in urban areas of low income countries. It is frequently observed that in such areas, the highest unemployment rate among young people, who also tend to be the most educated group. For educated women, it may even be rational to withdraw for as short period of time rather than take a low paying and low status job which may have unfavorable effect on their subsequent earnings potential.

In a comparative study of five developing countries, Smock (1981) examined some outcomes of educational attainment relative to women's overall activity rate and participation in modern sector. The result of his study does not support the general view of educational attainment positively related with LFP. However, Smock (1981) explained that in some cases there is only a little or none linear relationship between education and

Women are expected to be good mothers and homemakers while men are expected to be the bread winner and the head of the household in patriarchal family structure (Blau *et al,* 1998). This traditional division of work overburdens women with the unpaid reproductive tasks i.e. cleaning cooking, caring for child and elders etc, that women are usually restrained from labor market.

Thus FLFP of singles, divorced, widowed and separated (usually considered as unmarried) and married women are expected to differ. For instance, Youseef (1974) states that the influence of marital status and fertility are major conditions controlling the supply decision of women in labor market and his finding showed that single, widowed, and divorced women are most active in LFP than married women even if, the apparent willingness of married women to participate in LF stems from the desire to provide their families with higher standard of living and from other related forward pushes (Fosu ,1999)

The effects of a woman's social environment, as proxied by her residence, ethnicity and religion, on fertility and labor participation could be ambiguous. However, locality influence has to do with locality variation in the acceptability of women working, in occupational structure and in local labor market circumstance. Urban life is believed to be associated with various factors that help in reducing fertility and increasing a woman's participation rate. It must be noted, nevertheless, that a two-way effect could emerge from urbanization. There is a tendency that urbanization would increase women's opportunity cost of time by providing better earning possibilities as well as to reduce the cost of child raising quality by providing easier access to schooling.

In a sense, several studies explore the relation between participation and urbanization (Siegers and Zandanel, 1981, Evers and Van der Veen, 1983, and De Meester *et al,* 2007). All of these studies anticipate a positive relation between urbanization and labor market participation. There are two main arguments for this positive relation. Firstly, highly urbanized areas tend to have favorable labor conditions simply because there are more jobs available, which mean better opportunities of finding a job (Evers and Van der

Vaan, De Meester *et al*, 2007) and because firms and government offices which customarily employ a large number of women are predominantly located in highly urbanized areas (Siegers and Zandanel, 1981).

De Meester *et al* (2007) also mention the positive effect of supporting services in urban areas. Secondly, urbanization can be viewed as a substitute for the degree of emancipation (Evers and Van der Veen, 1983). Similarly, Siegers and Zandanel (1981) argue that the traditional opposition in society against occupational activities of women outside the home is probably less the more the region concerned is urbanized. Therefore, the overall effect of urban locality on female labor participation and fertility could be said to be more of an empirical issue.

The sectors of economy in which women were participating are frequently used to describe the labor market structure of the local market situations. Women in most developing countries are more likely to participate in the informal sector, where, they become family workers or self-employed. This occurs not only because it is easier for them to work while looking after children, but also because the fixed cost of employment in the informal sector is much lower than in the formal sector. Anker and Hein (1986) analyzed women's occupational distribution among six main nonagricultural occupations. Their findings suggest that in most Asian countries under that study, females dominate clerical and service occupations. In some countries, women are also engaged frequently in sales. However, the category 'sales worker' is very heterogeneous. Such studies, however, treat market employment as homogeneous while in most developing countries distinguishing between formal and informal sectors of the labor market is absolutely essential, especially regarding female labor participation.

Hill (1983) suggests that female labor participation behavior varies between the formal and informal sectors. For example, husbands' wages correlate negatively to women joining the formal sector but positively to women working for family business. Furthermore, having children fewer than six discourages women from working in the formal sector while encouraging women to work in family business. These findings

Vaan, De Meester *et al*, 2007) and because firms and government offices which customarily employ a large number of women are predominantly located in highly urbanized areas (Siegers and Zandanel, 1981).

De Meester *et al* (2007) also mention the positive effect of supporting services in urban areas. Secondly, urbanization can be viewed as a substitute for the degree of emancipation (Evers and Van der Veen, 1983). Similarly, Siegers and Zandanel (1981) argue that the traditional opposition in society against occupational activities of women outside the home is probably less the more the region concerned is urbanized. Therefore, the overall effect of urban locality on female labor participation and fertility could be said to be more of an empirical issue.

The sectors of economy in which women were participating are frequently used to describe the labor market structure of the local market situations. Women in most developing countries are more likely to participate in the informal sector, where, they become family workers or self-employed. This occurs not only because it is easier for them to work while looking after children, but also because the fixed cost of employment in the informal sector is much lower than in the formal sector. Anker and Hein (1986) analyzed women's occupational distribution among six main nonagricultural occupations. Their findings suggest that in most Asian countries under that study, females dominate clerical and service occupations. In some countries, women are also engaged frequently in sales. However, the category 'sales worker' is very heterogeneous. Such studies, however, treat market employment as homogeneous while in most developing countries distinguishing between formal and informal sectors of the labor market is absolutely essential, especially regarding female labor participation.

Hill (1983) suggests that female labor participation behavior varies between the formal and informal sectors. For example, husbands' wages correlate negatively to women joining the formal sector but positively to women working for family business. Furthermore, having children fewer than six discourages women from working in the formal sector while encouraging women to work in family business. These findings indicate that women's labor-participation behavior in developing countries varies across formal, informal and other relevant sectors. Thus, treating formal and informal sectors as a homogeneous labor market is very important to see the composition of female in such sectors of economy.

In spite of the above facts, women usually face the problem of unemployment more than their men counterparts due to barriers created especially by culture and misconception of employers about their capability and possibility of women leaving the job when they get married or have children (Bhatti, 2004). Thus, these invisible barriers are still preventing women to take a place at senior managerial administrative and legislative positions in most developing countries. Employers always use an excuse for not promoting women employment to administrative positions because of higher possibility of women leaving and reentering in the labor market in addition to the above socio-economic and demographic factors (Blau *et al*, 1998).

2.2 Urban Women in Labor Market of Amhara region

Amhara region constitutes the northern highland part of the country bordering Sudan in the North West, Tigray region in North East, Afar region in South East, Oromia region in South, and Benishangul-Gumuz region in the West. According to the 2007 census (CSA, 2007) Amhara region has a population of 17, 221, 976 of which 50 % were women.

Women and men have equal right and treated equally under the law in the regional state. Women have the same opportunities and rights for participating in economic, social, and political sphere of the country. Women and men have the same rights to work as recognized by the constitution of Federal Republic of Ethiopia as the basic human right which is provided in Article 41. In practice, the equality is not realized. In Amhara region women and men do not have the same opportunity in the labor market.

Women constitute 49% of the urban labor force of the region and their LFP was only 37% in 2005 (CSA, 2005) and in 2009 was 41% against 62 % of their men counterparts (CSA, 2009). Moreover women participation rate varies between major urban centers of the region. For instance, women participation rate for Gondar, Dessie and Bahardar in 2009 were 35, 37, and 46 % respectively. Women's share in employment is also low in the region where the share of women and men employment to total employment is 37% and 63%, respectively. A significant share of the employed in the service sector (70%) is higher than and only in this sector, the women's employment share is higher than those of men's (73.3% for women and 68.5% for men). Thus, women participation is concentrated in jobs, which are accepted as "female jobs", and they are in the jobs which are low productive, low paid and lower ranked. About 38% of them work as clerical and sales workers while 42% of men work in non agricultural production, equipment operators, assemblers and laborers.

In Amhara region 42% of the total employed are salaried workers, 44.3% are self employed and only 13.7% are unpaid family workers. Thus, out of total 44.3% share of self employed workers, women constitute only 43.3%. When the data from the 2009 UEUS analyzed separately for women and men for the region, the result indicates that paid employment is relatively scarce and self employment is dominant in the region for both sexes. About 42% of women and 44.5% of men, who were self employed had been dominantly engaged in informal economic activities and only 39.3% of the women employed in labor force are paid workers while salaried men constitutes 44% of the men labor force.

The 2007 population census indicated a slightly higher unemployment rate for women than men i.e. 17.3 for women 11.7 for men. Moreover, according to the 2009 UEUS women's unemployment rate is 3 times higher than that of men (8% for men and 23.4% for women). Similarly, the young women unemployment rate is on average 30% which is 3.5 times higher than those young men (8.6%) of the region. Women of the region usually face the problem of unemployment more than their men counterparts due to barriers created especially by tact socio-cultural and religious believes towards women and also misconception of employers about their capability and potential to take up higher managerial position. In 2009, women at higher position in public and private sectors constitute only 21% of the total of the region.

CHAPTER THREE

3. DATA AND METHODOLOGY

3.1 Source of Data and Description

The source of data for this study is the 2009 Urban Employment and Unemployment Survey (UEUS), conducted by CSA, Ethiopia, to provide statistical data on characteristics, size and distribution of the LF and status of economic activities of different subgroups of the urban population on annual basis.

In the 2009 UEUS a mixture of systematic two- and three-stage stratified cluster sampling design was used. All urban centers of the country were stratified into two groups: (1) the major urban centers which include all regional capitals and four other urban centers having populations higher than other urban centers such that from these 15 major urban centers of the first group, 331 sample enumeration Areas (EAs) and 11,130 sample households were selected. (2) The second group contains other urban centers. From this group, 82 sample urban centers, 270 sample EAs and 8,100 households were selected. The 2009 UEUS covered approximately 12, 119, 898 individuals, of which 52.2% were women.

In this study, the domains are the three major urban centers of Amhara region, namely Dessie, Gondar, and Bahardar towns. The 2009 UEUS covered a total of 2063 households and 7654 individuals in these major urban towns. From these 3014 women between ages 15 to 65 were identified for this study.

3.2 Variables in the Study

As it has been pointed out in the literature review the socio-demographic characteristics which were assumed to have significant effect are included in this study as explanatory variables/factors in determining the probability of female participation in labor force.

The explanatory variables/factors, which are considered in the analysis as potential determinants of women participation in labor force are: Educational attainment, Presence of children (number of children under 6 years of age old), Age of women, Marital status, and Relation to the Household Head. Age is included to capture the possible life cycle effect. Using these variables/factors in the three towns help us to gauge their effect independent of the effect of town of residence

In this study, the participation of women in labor force (LFP) is the response variable which is dichotomous taking a value 1 if the response of a woman was participating in the LF and zero otherwise. That is, if the woman response was participating in paid employment, engaged in self employment activities, unpaid family worker or a member of cooperative or member of board, the response is coded one otherwise zero. The descriptions of the explanatory variables/factors and the response variable with their code of their categories are presented in the following Table 3.1.

Table 3.1 Definition of variables/factors included in this study with their descriptions of their categories

No	Variables/Factors	code	Description
	The response variable		
1	Labor Force Participation (LFP)	1	If a woman is participating in the labor force
		0	otherwise
	Explanatory variables/factors		
2	Educational Status (EDUC_STATUS)		Dummy variables capturing women who are from education level
		1	Illiterate
		2	Primary or Junior school
		3	High school and above
3	Age(AGE)		Dummy variables capturing women of
		1	Age 15-22
		2	Age 23-33
		3	Age 34-44

3.3.1. Logistic Regression

Logistic regression can be used to predict a response variable on the basis of continuous and/or categorical independent variables with or without interaction terms and to estimate and determine the strength and the importance of the effects of these variables on the response. Logistic regression unlike ordinary regression does not assume linearity of the relationship between the response variable and the explanatory variables. Also it does not require the normality assumption as well as the assumption of homoscedasticity and in general has less stringent requirements. Of course, logistic regression analysis extends the techniques of multiple regression analysis to the research situations where the response variable is categorical.

In general, however, when the response variable is dichotomous, the ordinary least square regression technique provides parameter estimates that are inefficient, and heteroscedasticity error structure. As a result, tests of hypothesis and construction of confidence interval becomes inaccurate and misleading. Similarly, a linear probability model like the multiple logistic regression may generate predicted value that would spread out over an interval 0 and 1 and/or even negative values which violate the basic tenets of probability.

To alleviate these problems and to produce reliable outcomes, the most widely used qualitative response models are the logit and probit model. However, the logit model has the advantage over probit that the predicted probability could be arrived at easily and it is less sensitive to outliers and easy to correct bias. Moreover, the logistic regression model is a flexible model.

3.3.2. Logistic Regression Model

The logistic regression model can have an arbitrary number of explanatory variables/factors of either continuous, categorical or interaction terms. For simplicity consider a linear probability model

$$\pi(x) = \beta_0 + \beta x \quad \ldots(1)$$

where, $\pi(x)$= P(Y=1) is the probability (proportion) of cases for which y=1, β_0 is the intercept of the line, and β is the slope.

Thus, if we ignored the binary nature of Y and fitted the model as usual, the probability $\pi(x)$ may fall between 0 and 1 as the linear function takes values over the entire real line. However, for its extension with multiple predictors, difficulties often occur fitting this model because during fitting process the estimate of π(x) falls outside the $[0,1]$. Moreover, for binary responses, the constant variance condition that makes least squares estimator optimal is not satisfied. Since var(Y) = π(x)(1-π(x)), the variance depends on x through its influence on π(x) values (Collett,2003).

In general, the linear probability model has structural defects in various ways in estimating the probability in the range $[0,1]$. In order to avoid the problem of estimating the probability π(x) = P(Y=1) directly for binary response from the model (1) it is better to use the odds: the odds of success is defined to be the ratio of the probability of success to the probability of failure, i.e. $\frac{\pi(x)}{1-\pi(x)}$, which assumes any values from 0 to infinity. Beside this, when two sets of binary data are to be compared, a relative measure of the odds of a success in one set relative to that in the other is called the odds ratio.

Thus, the odd has form:

$$\frac{\pi(x)}{1-\pi(x)} = e^{\beta_0+\beta x} \quad \ldots (2)$$

From (2), non linear relationship between π(x) and x are often monotonic, with π(x) increasing continuously or decreasing continuously as x increases, which depicts an S-shaped curve. The curve with this shape derived from (2) has form.

3.3.4. Fitting the Logistic Regression Model

Analogous to multiple regressions Analysis, there are two important stages in the analysis of the data in logistic regression analysis. First, estimates for the unknown parameters in the model must be obtained and second, some determination must be made of how the model actually fits the observed data. In logistic regression analysis, the unknown parameters that must be estimated from the available data are the intercept β_0, and the logistic regression coefficient β_j. Because of the nature the model, maximum likelihood estimation technique is appropriate for estimating the logistic model parameters. In the context of logit model, maximum likelihood estimation principle involves by defining the likelihood, L of the sample data.

Suppose that the binary data of the y_i's taking value 1 or 0 for i = 1, 2,…, n observations. $\pi_i = P(y_i=1)$ is to be modeled as a linear combination of the values of p explanatory variables $x_1, x_2, \ldots, x_p$ that is indicated in equation (4) above. The observations y_i are from a binomial distribution with mean p_i, i.e. $y_i \sim \text{bin}(1, p_i)$. Then, in order to fit a linear logistic model for the given set of data, the $p+1$ unknown parameter $\boldsymbol{\beta}' = (\beta_0, \beta_1, \ldots, \beta_p)$ have first to be estimated. These estimates can readily be obtained using the method of maximum likelihood, and the likelihood function is given by:

$$l(\boldsymbol{\beta}) = \prod P(y_i / X_i, i = 1, \ldots, p) = \prod \left[\left(\frac{e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_i}}{1 + e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_i}} \right)^{y_i} \left(\frac{1}{1 + e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_i}} \right)^{1-y_i} \right]$$

where, y_i is 0/1 outcome of the i^{th} observation and $x_{i1}, \ldots, x_{ip}$ are the values of the predictor variable for the i^{th} observation based on a sample of n observations. This likelihood function depends on the unknown success probabilities π_i, which in turn depends on the $\boldsymbol{\beta}$'s through equation (5), and so the likelihood function can be regarded as a function of $\boldsymbol{\beta}$.

Our main objective is to obtain those values $\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p$ which maximize $l(\boldsymbol{\beta})$ or equivalently log $l(\boldsymbol{\beta})$. The logarithm of the likelihood function is:

$$\log l(\beta) = \sum [(y_i \log \pi_i + (1 - y_i)\log(1 - \pi_i))]$$

To find the value of **β**, we differentiate this log likelihood function with respect to the p+1 the unknown parameters are given by:

$$\frac{\partial \log l(\beta)}{\partial \beta_j}$$

where, $j = 0, 1, \ldots, p$.

Then evaluating these derivatives at **β** and equating them to zero gives a set of $p+1$ non linear equations in the unknown parameter β_j, that can only be solved numerically using an iterative algorism known as Fisher's method of scoring to obtain the maximum likelihood estimates of β. This algorism is implemented in most popular statistical packages like SPSS and SAS. For details of this algorithm (see Collett, 2003)

3.3.5. Model Building Strategies/Variable Selection

In building a model with many explanatory variables/factors, the first logical step is to select variables that result in the best model with the context of the scientific problem of interest. Thus, the model based on the selected variables should be complex enough to fit the data well and it should also be simple to interpret. Having this in mind, in this study, the variable selection process begins with a univariate analysis by studying the systematic relationship or association between a single predictor variable and the response variable.

Upon the compilation of univariate analysis, the predictor variables to be candidates for multiple logistic analysis will be selected with a condition that any variable whose univariate test has a p-value less than 0.25 was considered as a candidate for multiple model along with socio- demographic importance.

In multiple regression analysis, variable selection as well as model building process becomes harder as the number of predictors increases, because of the rapid increase in possible effects and interaction terms in the model. The importance of each variable included in multiple logistic model can be selected by making use of stepwise selection procedures, in which variables are selected either for inclusion or exclusion from the model in sequential manner based on statistical criteria only. The two most commonly used procedures for selecting a list of variables, which will result in joint impact on outcome variable, are forward selection and backward elimination procedures. This study employed a forward selection procedure. In forward selection, variables are added to the model one at a time. At each stage in the process, the variable added is the one that gives the largest decrease in the value of the deviance upon its inclusion. The process ends when the next candidate for inclusion in the model does not reduce the value of the deviance by more than a pre-specified amount. However, this is not the only sensible criteria for adding or dropping a term (the variable) in the model. The final decision for adding or dropping was made in accordance with prior tests and examination made on those variable.

3.3.6. Measure of Goodness of Fit

After fitting a model to a set of data, it is natural to inquire about the extent to which the fitted values of the response variable under the model are compared with the observed values. If the agreement between the observations and the corresponding fitted values is good, the model may be acceptable. If not the current form of the model will certainly not be acceptable and the model will need to be revised. This aspect of a model is referred to as test of goodness of fit.

In testing the hypothesis that the model fits the data against the alternative, the model does not fit the data well, there are a variety of statistical tests for determining the significance or goodness-of-fit. The most commonly used statistical approaches are Pearson's χ^2 statistic and the Likelihood-ratio statistic G^2, which are based on the comparison of the fitted and the observed values.

The Pearson's χ^2 statistic is given by:

$$\chi^2 = \sum\left[(observed - fitted)^2 \Big/ fitted\right] = \sum\left[\frac{\left(y_i - n_i \hat{p}_i\right)^2}{n_i \hat{p}_i(1-p_i)}\right]$$

where, n_i is the number of cases for a predictor $\boldsymbol{x}_i$, $\hat{p}_i$ is the fitted probability, and $\sum n_i = n$

The Likelihood-ratio statistic G^2 is:

$$G^2 = -2(L_M - L_S) = -2 \log \left(\frac{likelihood_M}{likelihood_S}\right)$$

where the likelihood ratio L_M is for the working model having few parameters and L_S is for the saturated model which has separate parameter for each observation (or has extra parameters) over the working model. Large values of χ^2 and G^2 suggests the lack of the model fit.

Another alternative approach for checking model fit is based on the Hosmer-Lemeshow test. The approach in Hosmer-Lemeshow test is by partitioning the data into 10 groups (g=10) and form observed and fitted values based on their estimated probabilities. This test uses Pearson's statistic to compare the observed and fitted counts for this partitions, which is given by:

$$C = \sum_{k=1}^{g=10} \frac{\left(O_k - n_k \bar{p}_k\right)^2}{n_k \bar{p}_k (1-\bar{p}_k)}$$

where, g is the number of groups, n is number of covariate patterns in the k^{th} group.

$O_k = \sum_{j=1}^{n_k} y_j$ the number of responses among n_k covariate patterns, and

$$\bar{p}_k = \sum_{j=1}^{n_k} \frac{m_j \hat{p}_j}{n_k}$$

where $\bar{p}_k$ is the average estimated probability and m_j is the number of observation for k^{th} group.

Calculated value of the Hosmer-Lemeshow test statistic C is greater than 0.05 or relatively small; it is an indication for the model adequately fit the data. Under the null hypothesis the statistic C has an approximate chi- square distribution with (g-2) degree of freedom (Hosmer and Lemeshow, 1989).

The Wald test is also another alternative test commonly used to test the significance of each individual parameter effect on the response. That is to test the null hypothesis that a particular logit effect or β_i has no effect on the response i.e. $\beta_i=0$ against $\beta_i \neq 0$. For binary response, the Wald statistic W is given as:

$$W = \frac{\hat{\beta}_i^{\,2}}{\operatorname{var}(\hat{\beta}_i)}$$

where, under the null hypothesis W has χ_{α}^{2} distribution with one degree of freedom.

3.3.7. Model Diagnostics

When a model is fitted to the observed values of a binary response variable, it is essential to check in detail that the extent to which the fitted model provides an appropriate description of the observed data is a vital aspect of modeling process. From a scientific perspective, the approach of comparing the working model to a more complex model is more useful than a global goodness-of-fit test (Agresti, 2000).

A large value of goodness-of-fit test statistic merely indicates some lack of fit, but do not provide an insight about its nature. Comparing a model to a more complex model, on the other hand, indicates whether lack of fit of particular type exists. In either of these cases when the fit is poor, diagnostic techniques can be used to measure the influence of individual observation on the model fit and also to highlights reasons for the inadequacy.

There are a number of reasons why a fitted model may not be adequate. The most common of these is that, the data may contain particular observations termed as outliers that are not well fitted by the model or observation termed influential values, which have undue effect of parameter estimation as well as conclusion to be drawn from the analysis (Collett, 2003).

3.3.7.1 Identification and Treatment of Outlying Observations

Observations that are surprisingly distant from the remaining observations in the sample are termed outlying values, or more briefly, outliers. Such values may occur as a result of measurement error, that is, error in reading, calculating or recording a numerical value, they may be due to an execution error, in which a faulty experimental procedure has been adopted, or they may be just on extreme manifestation of natural variability (Collett, 2003). Thus, to detect or identify such outlying observation, a number of diagnostic techniques have been proposed for use. The statistical techniques which are employed in this study for handling and treating outlier(s) are Standard Pearson residual and the Deviance residual.

Suppose that a linear logistic model is fitted to binary observations of the form y_i, $i = 1, 2, \ldots, n$ and the corresponding fitted value of y_i is $\hat{y}_i = n_i \hat{p}_i$. Then, the i^{th} raw residual is then $y_i - n_i \hat{p}_i$, and provides information about how well the model fits. However, using these raw residual is difficult to interpret and assess the model fit. In particular a large difference $y_i - n_i \hat{p}_i$ has a low precision (Collett, 2003). The precision of $\hat{y}_i$ is reflected in its standard error, given by:

$$\mathrm{se}(\hat{y}_i) = \sqrt{\hat{p}_i(1 - \hat{p}_i)}.$$

A low precision corresponds to a high standard error. The raw residuals can be made more comparable by dividing the residual by standard error se ($\hat{y}_i$), giving

$$e_i = \frac{y_i - n_i \hat{p}_i}{\sqrt{n_i \hat{p}_i (1 - \hat{p}_i)}}$$

These residuals are known as Pearson residuals and they measure the contribution that each observation makes to a statistic that is a summary measure of the goodness of the fitted logistic regression model. For this reason, they are intuitively appealing as measures of model adequacy.

A better procedure to identify outliers is to divide the raw residuals by their standard error,

se (y_i - $\hat{y}_i$) , is given by se(y_i - $\hat{y}_i$)= $\sqrt{\hat{v}_i (1 - h_i)}$, where $\hat{v}_i$= $\hat{p}_i$ (1- $\hat{p}_i$) and h_i is the i^{th} diagonal element of the $n \times n$ matrix H = $V^{1/2}X(X'VX)^{-1}X'V^{1/2}$ in this expression for H, V= Diag ($\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n$) is the diagonal matrix of weights, the estimated variance of y_i used in fitting the model, and X is the $n \times p$ design matrix, where p is the number of parameters in the model. The resulting standardized residuals are

$$r_{p_i} = \frac{y_i - n_i \hat{p}_i}{\sqrt{n_i \hat{p}_i (1 - h_i))}} .$$

These residuals are simply the Pearson residuals, e_i, divided by $\sqrt{1 - h_i}$ and are therefore known as standardized Pearson residuals. These residuals are used to measure the extent to which each observation deviates from its fitted value under the assumed model. Consequently, an observation that leads to abnormally large residual is an outlier.

Another alternative residual uses components of G^2 fit statistic is the deviance residuals. The deviance residual for observation i is

$$\text{sign} (y_i - n_i \hat{p}_i) \sqrt{d_i}$$

where, $d_i = -2\ \{y_i \log \hat{p}_i + (1-y_i) \log (1-\hat{p}_i)\}$.

Thus, absolute values of such residuals larger than roughly 3 provide evidence for possible existence of an outlier(s) (Agresti, 2000 and Collett, 2003).

Whenever outliers are identified, their effect on the results of the analysis can be assessed by re-analyzing the data after omitting them. If essentially the same inferences are drawn from the data, both with or without the outlier(s), one needs not be too concerned about their presence. On the other hand, if the outliers do not affect model-based inferences, the decision on how to regard them will be crucial, but the decision should not be made on statistical ground alone, subjective decision may be necessary how to treat them (to include or omit them) (Collett, 2003).

3.3.7.2 Identification and Treatment of Influential Observations

An observation is said to be influential if its omission from the data set results in substantial changes to certain aspects of the fit of the linear logistic regression model. In most cases, it is more important to pay attention to outliers that are influential than those that are not. Although outliers may also be influential observations, an influential observation need not necessarily be an outlier. In particular, an influential observation that is not an outlier will occur when the observation distorts the form of the fitted model to such an extent that the observation itself has a small residual. Thus, the presence of an influential observation cannot necessarily be detected from direct examination of the residuals, and so additional diagnostic techniques are required (Agresti, 2000 and Collett, 2003). To do so, some of statistical techniques which will be employed in this study for handling and treating outlier(s) and/or influential observations are leverage statistic, Cook's distance, and Dfbeta's.

The leverage statistic, h is used to identify cases which influence the logistic regression model more than others. The leverage statistic varies from 0 (no influence on the model) to 1 (completely determines the model). The leverage value of any given observation can be compared to 2p/n, which is twice of the diagonal elements of the H matrix defined earlier, where p is the number of parameters in the model. On the other hand, one can also use to detect those outliers which are influential possibly by plotting the deviance residuals with the leverage value h_i.

To examine how the i^{th} observation affects the set of parameter estimates, the vector of parameter estimates for the full data set, $\hat{\beta}$ is compared with the vector of estimates $\hat{\beta}_{(i)}$, obtained when the same model is fitted to the data excluding the i^{th} observation. Thus, the most useful statistic for measuring the extent to which the set of parameter estimates $\hat{\beta}$ is affected by the exclusion of the i^{th} observation is by using Cook's distance (C_i), given by

$$C_i = \frac{h_i r^2{}_{p_i}}{p(1-h_i)};$$

where, p is the number of unknown parameters in the model $r^2{}_{p_i}$ is the square of the standardized Pearson residual and h_i is the i^{th} observation leverage value. Hence an observation for which its Cook's distance greater than one is considered as influential.

Cook's distance measures the overall effect of the i^{th} observation on the parameter estimate. In certain situations, such as when some of the terms in the model are particular interest, one may want to examine the influence that each observation has on the j^{th} parameter estimate $\hat{\beta}_j$, when the model includes a constant term. The statistic which is used in this situation is Dfbeta's given by:

$$Defbetas\Delta_i \hat{\beta}_j = \frac{\hat{\beta}_i - \hat{\beta}_{j(i)}}{\sqrt{S_{(i)}{}^2 - D_{jj}}}$$

Where, $D = (X'\hat{V}X)^{-1}$ is the Variance-Covariance matrix of the parameter estimates, D_{jj} is the j^{th} diagonal element of D and x_i is the vector of explanatory variables for i^{th}

observation (for detail see Collett, 2003). Dfbeta greater than one suggest that the corresponding observation influences the relevant parameter estimate.

3.3.7.3 Multicollinearity

Multicollinearity in logistic regression is a result of strong inter-correlations among the predictor variables. Inter-correlations among the predictor variables by themselves need not necessarily cause any problems in inference. Whether or not this is a problem will depend on the magnitude of the error variance and the variances of the predictor variables. Multicollinearity may be induced due to poor sampling method, misspecification and over fitting of a model as well as improper use of dummy variables (Montgomery and Peck, 1992).

Some of the statistical techniques that have been proposed for detecting multicollinearity among categorical predictor variables are: inspecting off-diagonal elements of the Kendall's tau bivariate correlation matrix of the explanatory variables, variance inflation factors (VIFs), tolerance and condition numbers or indices. Examining the Kendall's tau bivaraite correlations between the predictors (say x_i and x_j for $i \neq j$) and looking for "large" values of r_{ij} that is 0.8 and larger could help detecting high multicollinearity (Shoukri and Edge, 1996). The problem, however, with this method is that one predictor may be a linear combination of several predictors, and yet not be highly correlated with any one of them. Variance inflation factors (VIFs) are probably superior in detecting multicollinearity to the analysis based on bivaraite correlations and defined as

$$\text{VIF}\ (\hat{\beta}_j) = \frac{1}{1 - R_j^{\,2}}$$

where, R_j^2 is the coefficient of determination obtained when x_j is regressed on the remaining p-1 predictors. The VIF for each term in the model measures the combined effect of the dependencies among the predictors on the variance of that term. Commonly when VIFs exceed 10 this is an indication that the associated regression coefficients are poorly estimated because of multicollinearity.

CHAPTER FOUR

4. DATA ANALYSIS AND DISCUSSION

Female participation in labor force (FLFP) is a dichotomous response variable with value 1, if a woman was participating in labor force or 0 otherwise at the point of time of the survey in 2009. In this study, logistic regression model were fitted to identify and examine the potential determinant factors influencing LFP of women in three towns of Amahra region, namely Dessie, Gondar, and Bahardar. Having identified potential determinant factors using univariate logistic regression we then applied multiple logistic regression models to find out the overall effect of identified factors on FLFP of the three major towns of the region. In the following subsequent sections we discuss the influence of the factors on LFP.

4.1. Summary Statistic for the Sample Data

During the specified period, a total of 3014 women were interviewed in the three major towns of the region in UEUS. Out of which 948, 1065, and 1001women were from Bahardar, Gondar and Dessie town respectively. The percentage distributions of FLFP for different Age groups, marital status, Presence of children, Educational status, and Relation to the head, with respect to each town of residence were calculated. These absolute figures and percentages are shown in Table 4.1.

Socio-Demographic Attributes and LFP related to Town of Residence

The table shows that labor force participation rate of an individual woman differs from town to town. LF participation of women is: in Bahardar town (50.9%), in Dessie (42.0%), and in Gondar town (38.0%). Table 4.1 gives information about the various personal and household attributes that contribute for the difference in FLFP among these towns.

FLFP rates appear to vary among all age groups and depict somehow an inverted U-shaped age profile. Higher LFP of women in the age groups of 23-33 and 34-44 are found to be 67.5% and 70.3%, for Bahardar and Dessie towns respectively, while it is found to be only 50.5% for Gondar town. In contrast to age profile, LFP related to educational status for Gondar and Bahardar shows somehow U-shaped curve, indicating that higher LFP was for women in the illiterate group and in the most educated group (high school and above), while lower participation was for those women having elementary or junior school education. For Dessie town proportion of women LFP shows decreasing trend as the level of women education increases.

Table 4.1 also shows the FLFP profile by marital status. In this regard the table depicts there is no significant difference among married and unmarried (singles, widowed or divorced) women of Gondar (35.9% and 39.4%) and Bahardar (49.7% and 51.6%), and also true for women of Dessie (44.0% and 40.7%). In terms of presence of children, women from the HH with no child and at least one child at home have similar proportion in LFP in Dessie town (41.4% and 43.0%). Contrary to this, LFP of a woman from the HH with no child or at least one child less than 6 years old at home were different in Gondar (40.2% for no child and 34.5% for at least one child) and also in Bahardar town (for no child, 53.2% and for at least one child, 46.6 %). These figures revealed that, a high proportion of women LFP is observed for those women from the household with no child at home in both Gondar and Bahardar towns.

derived from the estimated logistic probabilities) by the observed category values of LFP. Thus, a good model is one that minimizes the misclassification.

Table 4.11 Classification Table for Gondar Town [a]

Observed		Predicted		
		Labor Force Participation (LFP)		Percentage Correct
		Unemployed	Employed	
Labor Force Participation	Unemployed	549	109	83.4
	Employed	240	163	40.4
Overall Percentage				67.1

a The cut value is .500

Table 4.12 Classification Table for Dessie Town [a]

Observed		Predicted		
		Labor Force Participation (LFP)		Percentage Correct
		Unemployed	Employed	
Labor Force Participation	Unemployed	461	118	79.6
	Employed	227	192	45.8
Overall Percentage				65.4

a The cut value is .500

Table 4.13 Classification Table for Bahardar Town [a]

Observed		Predicted		
		Labor Force Participation(LFP)		Percentage Correct
		Unemployed	Employed	
labor force participation	Unemployed	305	159	65.7
	Employed	150	333	68.9
Overall Percentage				67.4

a The cut value is .500

From Table 4.11 the overall predictive efficiency of the fitted model for Gondar town is about 67.1 % meaning that 67.1 % of them are correctly predicted by the model. In terms of sensitivity and specificity, 83.4% of unemployed women and 40.4% of employed women are correctly predicted. The classification Table 4.12 for Dessie town reveals that 65.4% of women are correctly predicted by the model. Thus, 79.6% and 45.8% of unemployed and employed women respectively are correctly classified in their respective categories. From Table 4.13 for Bahardar town, 67.4% of women are correctly predicted (classified) by the fitted model while 65.7% of unemployed and 68.9% employed women are correctly classified or predicted.

Table 4.20 the Hosmer and Lemeshow Test for model fitted for Dessie Town

Step	Chi-square	df	Sig.
1	13.350	8	.100

Table 4.21 the Hosmer and Lemeshow Test for model fitted for Bahardar Town

Step	Chi-square	df	Sig.
1	22.225	8	.005

Despite this fact, it is already illustrated in the previous test result that, for Bahardar and Gondar towns the fitted models are adequate when compared to the model fitted for Dessie town. However, we cannot take the Hosmer-Lemeshow test result particularly for Bahardar and Gondar town because the numerical values of the Hosmer-Lemeshow test statistic depend on the number of groups chosen (used) the number of observations within the groups, and on how sets of observations with a common fitted probability are split. Thus, P-value of the test should not be interpreted rigidly, and is best taken to be an informal guide as to the adequacy of the model (Collett 2003, p88).

Furthermore, once a model has been fitted to the observed values of a binary response variable, it is essential to check the validity of the fitted model. Indeed, the global goodness of fit test was used to examine the extent to which the fitted models describe the observed data. The results indicate that the fitted models did fit the data well. But further diagnostic test are conducted for the presence of some observations that are not well fitted by the model which might affect the fit of the model as well as the parameter estimation. Therefore, examining the adequacy of the fitted models through model diagnostic techniques can help us to detect observations which could affect fit of the model and accordingly to take appropriate diagnostic measures.

Thus, the adequacies of the fitted models were checked for the presence of possible outliers, influential observations, and multicollinearity. The test results for detection of outliers and influential values are presented in the Appendix for each town. The results in Tables A3.1, A3.2, and A3.3 show that absolute value of the standardized Pearson residuals and the deviance value are both less than 3 in absolute value, implying that there are no possible outliers for the data of each town. For each town the observation leverage values are all less than unity, implying that there are no potential influential observations that affect the fit of the model. The DFBETA's for individual model parameters including their constant terms are less than unity. This implies that there are no specific influential observations on the estimate of a coefficient a particular predictor variable. Moreover, for the three towns the Cook's distance is less than one implying that no observations have influence on the estimated vector of regression coefficients.

Multicollinearity diagnostic test was also done using the condition number (indices), tolerance, VIFs (Variance Inflation Factors) and Kendell's tau_b correlation matrix. The results for the three towns are presented in Appendix Tables A1.1, A1.2, A1.3, A2.1, A2.2, and A2.3. All of the tolerance values are close to 1. The correlation between any two predictors is not greater than 0.8. These results imply that multicollinearity is not a serious problem. The condition numbers were found to be between 27.86 and 38. These results suggest that, there is weak multicollinearity, which does not affect the model fit. In summary, all of the above test results reveal that all of the fitted models of Table 4.3, 4.4, and 4.5 were adequately fit their data well.

4.5 Interpretation and Discussion of the Results

The results of the multiple logistic regression analysis agree with the results obtained from summary statistics and univariate analysis. Of course, in the univariate analysis, marital status was not significantly associated with FLFP for the three towns but it was found to be significant in the multiple analysis except for Gondar at 0.1 level of significance. On the other hand, educational status for Dessie town is found to be significantly associated with LFP at the significance level of 0.25 but not in the multiple analysis for both Dessie and Bahardar towns. In all three major towns of Amhara region only Age and Relation to the HH head were found to be significant factors in determining women LFP. However, the effect of remaining variables depends on a woman town of residence such that only presence of children and educational status variables for Dessie town, presence of children for Gondar town and Educational status for Bahardar town are not found to be significant in determining FLFP.

alleviate the limiting effect of child care on women's labor market participation especially in Bahardar town. Of course, such efforts could contribute in reducing the negative effect of marriage.

- In conclusion, it is essential to say that women's labor has remained a neglected area, although some efforts have been made recently. Great deal of research needs to be done at national level to generate necessary information to determine other factors influencing women's labor supply decisions.

REFERENCES

Anker. R. and Hein, C. (1986), *Sex Inequalities in Urban Employment in the Third World*. NewYork: St.Martin's Press.

Agresti, A. (2000), *An Introduction to Categorical data Analysis (2nd Ed.),* John Wiley and Sons, Inc., New York.

Bhatti, F. (2004), "Gender (In) Equality in North Cyprus Labor Market: Benchmarking Analysis with European Union Member and Candidates", Review of Social Economics and Social Studies **5/6**, 200-256.

Bindary, A., Colin, B. B., and Hollingsworth, T. H.(1973), "Urban-Rural Differences in the Relationship between Women's Employment and Fertility: A Preliminary Study", *Journal of Biosocial Science* 5:159-67.

Blau, F.D., Ferber, M.A., and Winkler, A.E. (1998), *The Economics of Women, Men, and Work* (3rd Ed.), London: Prentice Hall Inc.

Bowen, W. G. and T. A. Finegan (1969), *The Economics of Labor Force participation*. Princeton, NJ: Princeton University Press.

Brewster, K.L. and Rindfuss, R.R. (2000), "Fertility and women's employment in industrialized nations", *Annual Review of Sociology*, **26**, 271–296.

Cramer, J.C. (1980), "Fertility and female employment: Problems of causal direction". *American Sociological Review*, **45**(2), 167–190.

Collett, D. (2003), *Modelling Binary Data* (2nd Ed). Chapman and Hall/CRC. London.

CSA, Central Statistical Agency (2007), "Summary and Statistical Report of the 2007 Population and Housing Census". December 2008, Addis Ababa.

CSA, Central Statistical Agency, (2009), "Report on the 2009 Urban Employment and Unemployment Survey". *Statistical Bulletin* **469,** December 2009, Addis Ababa.

CSA, Central Statistical Agency, (2005), "Report on the 2005 National Labor Force Survey", *Statistical Bulletin* **365,** May 2006, Addis Ababa.

De Meester, E., Mulder, C.H. and Droogleever F. J. (2007), "Time spent in paid work by women and men in urban and less urban contexts in The Netherlands", *TESG*, **98** (5), 585-602.

Elder, S. and Johnson, L. J. (1999), "Sex-specific labor market indicators: what they show", International Labor Review", **138**(4), 447-464.

Evers, G. and Veen, A. v. d. (1983), "A simultaneous model of regional labor supply incorporating labor force participation, commuting and migration", *Socio-Econ.Plan. Sci.*, **17** (5-6), 239-250.

Fitzenberger, B., Schnabel, R., and Wunderlich, G. (2004), "The gender gap in labor market participation and employment: A cohort analysis for West Germany", *Journal of population economics*, **17**(1), 83-116.

Fosu, A.K. (1999). "Cost of living and labor force participation: Married women in urban labor markets", *Journal of Labor Research*, **XX**(2): 219–32.

Goldstein, S. (1972), "The Influence of Labor Force Participation and Education of Fertility in Thailand", *Population Studies* **26**:419-36.

Hill, C.R. &. Stanfford, F.P. (1985), *Lifetime fertility, child care, and labour supply*. In F.T. Juster & F.P. Stafford (Eds). *Time, Goods, and Well-being,* University of Michigan.191-224.

Hill, M.A. (1983), "Female labor force participation in developing and developed countries: Consideration of the informal sector", *The Review of Economics and Statistics*, **65**(3), 459–468.

Hosmer, D. and S. Lemeshow (1989), *Applied Logistic Regression,* John Wiley and Sons,Inc., New York.

Iacovou, M. (2001). "Fertility and female labor supply", *Institute of Social and Economic Research Working Paper No. 19*, University of Essex, Colchester.

Isvan, N. A. (1991), "Productive and Reproductive Decisions in Turkey: The Role of Domestic Bargaining", *Journal of Marriage and the Family* 53:1057-70.

Mackellar, L. and R. Bird (1997), "Global population ageing, social security and economic growth: Some results from a 2-regional model", Paper presented to Project LINK meeting. New York.

Mason, K.O. and V.T. Palan (1981), "Female Employment and Fertility in Peninsular Malaysia: The Maternal Role Incompatibility Hypothesis Reconsidered",*Demography* **18:**549-75.

Montgomery, D.C., and Peck, E. A. (1992), *Introduction to Linear Regression Analysis* (2nd Ed) ,John Wiley & sons, Inc.

Murray, P. and Syed, J. (2007), "Observations through gendered lenses: experiences of managerial women", In E. Davis & V. Pratt (Eds), *Making the Link: Affirmative Action and Employment Relations*. Sydney: CCH Australia Limited, 18: 38-44.

Newman, D. M. (1997). *Sociology: Exploring the Architecture of Everyday Life* (2nd Ed) Thousand Oaks, CA: Pine Forge Press.

Siegers, J. J. and Zandanel, R.(1981), "A Simultaneous Analysis of the Labor Force Participation of Married Women and the Presence of Young Children in the Family", *The economic review*, **129** (3), 382.

Sheehan, G. and Standing, G. (1978), "Economic activity of women in Nigeria" , in G. Standing and G. Sheehan (eds.), *Labor Force Participation in Low-Income Countries* (Geneva, ILO, 1978), 129-136.

Shoukri, M. M. and V.L. Edge(1996), *Statistical methods for health Sciences,* CRC press, Boca Raton.

Standing,G. (1978), "Labor Force Participation and Development", Geneva: International labor Office.

Smock, A.C. (1981), *Women's Education in Developing Countries: Opportunities and Outcomes*, Praeger Publishers, New York.

UNDF, United Nations Development Assistance Framework in Ethiopia (2001), "Contributing Towards Reducing Absolute Poverty in Ethiopia" The United Nation Country Team, Addis Ababa, Ethiopia.

UNDP, United Nations Development Fund (2003), *Human Development Indicators 2003*. New York: Oxford University Press.

Vlasblom, J. D. and Schippers, J. J., (2004),"Increases in female labour force participation in Europe: similarities and differences", *European Journal of Population* **20**(4), 375-392.

Weller, R.H. (1977), "Wife's employment and cumulative size in the United States, 1970 and 1960". *Demography*, **14**(1), 43–65.

Youssef, N.H. (1974), *Women and Work in Developing Societies*, University of California Press.

APPENDIX

Table A1.0 Categorical Variables Coding used by SPSS software

		Parameter coding		
		(1)	(2)	(3)
AGE	15-22	1	0	0
	23 - 33	0	1	0
	34 - 44	0	0	1
	45 - 65	0	0	0
RELATION	Head	1	0	
	spouse	0	1	
	Other family memeber	0	0	
TOWN_RESIDANCE	Gondar	1	0	
	Dessie	0	1	
	Baherdar	0	0	
EDUC_STATUS	Iliterate	1	0	
	Elementay and junior school	0	1	
	High school and above	0	0	
MART_STATUS	Never married, Single,divorced or widowed	1		
	Married	0		
PRES_CHILD	no children at home	1		
	at least one children	0		

DECLARATION

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

Name: FEKADU MULUYE BELAY

Signature: ....

Addis Ababa University

Collage of Natural Science

Department of Statistics

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

Prof. Eshetu Wencheko ....

Professor Eshetu Wencheko

Date: March 1, 2011