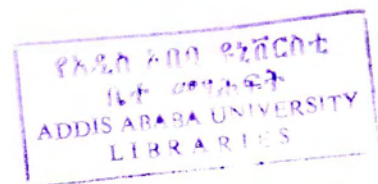


186

**FACTORS INFLUENCIUNG ENGAGEMENT
IN NON-FARM ACTIVITIES
THE CASE OF AMHARA REGION**

By

Abebaw Getachew



**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**

**JUNE, 2000
ADDIS ABABA**

Addis Ababa University
School of Graduate Studies

Title of Research Factors Influencing Engagement in Non-Farm Activities
The Case of Amhara Region
Name of Candidate Abebaw Getachew
Department Statistics
Faculty Natural Science

Approved by

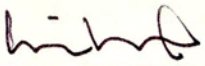
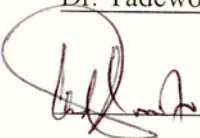
Advisor and Chairman	Examiner	Examiner
Name <u>Dr. Eshetu Wencheke</u>	<u>Dr. Tadewos Koroto</u>	<u>Dr. Butte Gotu</u>
Sign. <u></u>	<u></u>	<u></u>
		Date <u>19 June 2000</u>

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS.....	i
ABSTRACT.....	ii
CHAPTER I. INTRODUCTION	1
1.1 GENERAL	1
1.2 BACKGROUND OF THE AMHARA REGION.....	4
CHAPTER II. OBJECTIVE AND METHODOLOGY	7
2.1 OBJECTIVE.....	7
2.2 DATA SOURCE.....	8
2.3 VARIABLES OF THE STUDY	9
2.3.1 RESPONSE VARIABLE	10
2.3.2 PREDICTOR VARIABLES.....	10
2.4 METHODOLOGY.....	14
2.4.1 COMPARATIVE REVIEW OF SOME QUALITATIVE RESPONSE MODELS (QRM).....	14
2.4.2 CHOICE OF MODEL.....	17
2.4.3 PARAMETER ESTIMATION.....	18
2.4.4 VARIABLE SELECTION AND GOODNESS- OF - FIT ASSESSMENT OF THE MODEL.....	21
2.4.5 MODEL ADEQUACY.....	25
2.4.5.1 DETECTION AND TREATMENT OF OUTLIERS.....	25
2.4.5.2 INFLUENCE DIAGNOSTICS	26
2.4.5.3 MULTICOLLINEARITY	27
CHAPTER III. LITERATURE REVIEW.....	29
3.1 OVERVIEW OF STUDIES ON RNFA.....	29
3.1.1 THE ROLE AND TREND OF RNFA.....	29
3.1.2 DETERMINANTS OF RNFA	33

3.2 COMMENTS ON STUDIES ON RNFAS.....	36
CHAPTER IV RESULTS AND DISCUSSION.....	39
4.1 UNIVARIATE AND MULTIVARIATE RESULTS.....	39
4.1.1 UNIVARIATE RESULTS.....	39
4.1.2 MULTIVARIATE RESULTS.....	43
4.2 DISCUSSION.....	47
4.3 IDENTIFICATION OF WEALTH GROUPS IN RNFAS.....	53
CHAPTER V. CONCLUSION AND RECOMMENDATION	59
REFERENCES.....	63
APPENDICES.....	68
APPENDIX I. CLASSIFICATION TABLE.....	68
APPENDIX II. GOODNESS OF FIT SUMMARY STATISTICS	68
APPENDIX III-1. HALF-NORMAL PLOTS OF STANDARDISED RESIDUALS.....	69
APPENDIX III-2 CASE-WISE PLOTS OF RESIDUALS.....	70

Acknowledgments

This study is the result of the encouragement and support of a number of people. First and foremost, I am highly indebted to my advisor **Dr. Eshetu Wencheke** for his useful advice and editing the thesis timely despite the pressure of his regular work.

My thank also goes to the Amhara National Regional State particularly to the late **Ato Mesfin Kassa**, who was head of the Labour and Social Affairs Bureau of the region, for giving me the opportunity to pursue this postgraduate programme. The Ministry of Labour and Social Affairs is thanked for availing me the data for this study.

My grateful thanks are also due to **Dr. Mulat Demeke**, Associate professor in the faculty of Business and Economics in the Addis Ababa University, for editing this thesis and providing me very recent publications in the area. I would also thank him for his friendly suggestions, valuable comments and his hospitality. **Dr. Tegegne Gebre Egziabher**, Director of Institute of Development Research of the Addis Ababa Univerasity, is acknowledged for providing me his personal reference materials. I also want to thank ILO for allowing me to use their library facilities and appreciate the excellent cooperation of the staff of the organisation.

I am indebted to my intimate friend **Ato Asnakew Assefa**. His unreserved support, encouragement, friendly criticism coupled with his welcoming hospitality enabled me to complete this programme successfully. The encouragement of my friend **Nibretie Gobezie** was also invaluable.

I am particularly indebted to my wife **Muluken Asres** for her continuous encouragement, assistance and patience that gave me the courage to finish the two years programme with no difficulty.

Finally, many thanks are due to my brother **Ayele Berkie** for paving me the way to the present achievement and the support he has given me throughout my school life.

ABSTRACT

Some rural households of the Amhara region are counteracting the falling in agricultural output through combining agriculture with one or more non-farm activities. However, these activities, besides their low return, are influenced by a number of factors.

This paper delves on identifying the important demographic, endowment and related factors (predictors) of engagement in non-agricultural activities by employing the logistic regression model.

The results of the study suggest that the demographic variables sex, age, religion and marital status of heads, and household size; the endowment and related variables agricultural yield, and food balance are some of the important predictors of involvement in non-farm activities in the Amhara region.

It is found out that female-headed households are more likely involved in non-agricultural activities compared with male-headed ones. It is also found out that the overall trend of these activities decreases the older the head.

The results further revealed that households whose heads are followers of the Orthodox Church are less likely involved in income source diversification when compared with households headed by followers of other religions. As regards family size, rural households endowed with larger family size tend to diversify their income sources compared with those having smaller members.

The result of the study also indicated that households who are food insecure and who produced insufficient agricultural output are likely to engage in non-agricultural activities compared with those relatively better-off in their agricultural return.

It is recommended that the non-farm sector should be given due attention in development research. Moreover, the scope of credit schemes has to be expanded to enhance the productivity of non-agricultural activities in the Amhara region.

INTRODUCTION

1.1 General

The largest proportion of Africa's population is engaged in traditional agriculture. Its return is very low and food security is a major problem.

Ethiopia's agriculture is not sufficiently feeding the ever-increasing population of the country. A total of over 44 million people were affected by repeated drought in the former Gojjam, Gondar and Wollo provinces in 1978-88 (Relief and Rehabilitation Commission of Ethiopia (RRC), 1994). The production of major crops (*meher* season) showed a decline of 10.04 per cent in 1996/7 to 1997/8 and a reduction of about 14 % was recorded in the Amhara region in the same period (Central Statistical Authority (CSA), 1998). Traditional methods of farming, massive land degradation, bad weather conditions, and misguiding policies of the past governments are the main causes amongst others (Tegegne, 1996; Mulat and Teferi, 1996).

Evidences show that agriculture has failed to support even the subsistence livelihood of the rural population of the Amhara region despite measures taken to ameliorate the problem. It is documented, for example, that food insufficiency has prevailed since long years ago in North Shoa, one of the administrative zones of the Amhara Regional State (Mulat and Teferi, 1996). More than 65 percent of the households in this area were found net buyers of grain. The decline of livestock ownership aggravated the already deep-rooted problem of the area. As in any problematic parts of Africa, households of this impoverished part of Amhara region have been trying to circumvent food self-insufficiency through diversifying their income into different types of non-agricultural activities. It was pointed out that out of the total annual mean

income of Birr 1,235.72 derived in the area, the largest share (59.5 %) was accounted by non-farm income. The result exemplifies the contribution of non-agricultural income and, at the same time, the deterioration of the mainstay of the rural residents. Considering the overall situation of the labourious and risky traditional farming, Ethiopia has been trying to identify and encourage non-farm activities since many years.

A similar situation has been observed in other parts of Africa. Jambiya (1998), based on a survey on West Usambara Mountains and Lushoto Villages of Tanzania, found out that farm plots per household are decreasing due to the growing number of households. Agricultural returns have declined or stagnated and food insecurity and inadequacy prevailed in the study areas mainly due to land pressure, declining soil fertility, and other natural calamities. As it was described by Mung'ong'o (1998), the average agricultural revenue per cultivated land decreased by 22 percent and the annual mean household agricultural income had declined by 71 percent for the period 1972-92 in another part of the country called Njombe.

According to Mung'ong'o (1998) households dependent only on agriculture appeared to be economically less well-off and the increasingly less promising nature of agriculture forced many of the inhabitants to combine farming and non-agricultural activities as a survival strategy to meet the increasing number of basic requirements of households.

Sahelian households (Niger) are counteracting the falling agriculture by diversifying their income strategies through a variety of non-farm activities: employment on other local farms, migration to find employment, self-employment in a rural non-agricultural activity, employment within others' non-agricultural activities (Møller, 1998).

In rural Hausaland (Nigeria) the proportion of medium-scale and small-scale farmers suffering from grain deficits had increased by 16 and 17 percent respectively between 1980-90 and

1992-3 (Meagher and Mustapha, 1997). Similar problems were observed in hinterlands of Gusau and Sokoto towns of the country. Iliya and Swindell (1997) revealed that changes in agriculture of the areas had been related to the waxing and waning of non-farm incomes.

In Kenya, it has been widely recognised that confining to traditional agriculture is no longer a feasible livelihood strategy for most farm households (Foeken, 1997). Furthermore, it was believed that poverty alleviation is possible if sufficient off-farm opportunities are available.

Due to its problems over years traditional farming has been systematically replaced by modern farming and rural non-farm activities in Zimbabwe (Gaidzanwa, 1997). It was described that de-agrarianisation was perceived to be 'the shedding of the agricultural population to the towns and to the rural non-farm sector'.

Ibrahim (1997) described that the capacity of Sudanese agricultural sector is not sustaining the populace in most parts of the country. He mentioned that the overall share of rainfed agriculture's income from the total household's income in a region called, Darfur in 1991/2 was about 22.8 percent. Furthermore, about 54-95 percent of the inhabitants of the area experienced food shortage in 1984/5-1992/3 and the proportion of agriculture's contribution in household's food supplies was less than 30 percent in 1992/3. Petty trade, firewood and grass collection, and handicraft production (pottery, blacksmithery, mats and baskets, and carpets) were being considered as coping strategies in this area.

African rural households are practicing different non-farm activities to supplement traditional agriculture. Though these activities may depend on the type of household's livelihood and experience, they have been playing different roles: as a defense against famine and as social security mechanism - for the majority poor; as a means of supply of goods and services which would be beyond reach - for the majority medium and some poor class households; and as a

take-off for an accumulation and boosting agricultural production - for the wealthy (Møller, 1998; Bryceson, 1997). Their role in narrowing the gap of wealth differences of rural households is also undeniably high.

The different studies, conducted under different geo-economic situations, reveal that the share of traditional agriculture in Africa's economy is declining. Combining traditional agriculture with one or more non-agricultural activities is being considered as a survival strategy in most rural areas (Bryceson, 1997).

1.2 Background Information on the Amhara Region

Ethiopia is divided into 10 regional states and an administrative council. One of its states is the Amhara region.

The Amhara National Regional State (ANRS), which is situated between 8°45' N to 13° 45' N latitude and 35° 46' E to 40° 25' E longitude, covers nearly 15 percent of the area of the country (approximately 152025.24 km²). Its population as recorded in Oct. 1994 was about 13,834,297 (more than one-fifth of the country's population). The annual population growth rate at that time was about 2.86. The population density is nearly 91 persons/km², ranging from 151 in west Gojjam to 39 in Wag Himra (CSA, 1998). It embraces the majority of frequently drought affected and some of the most over-cropped parts of the country (Tegegne, 1995).

Its rural sedentary population, which was more than a fifth of the nation's rural population, accounted for nearly 91 percent (50.7 % males and 49.3 % females) of the region's population. Its total fertility rate is about 4.5, and family size about 4.5 (CSA, 1998). The part of the population of utmost 14 years of age (the minimum age for employment in the country) and

those above the age of 64 comprised 44.7 and 3.9 percent respectively of the people of ANRS, signifying a high dependency ratio in rural areas.

More than 68 percent of rural heads of households cannot read and write (Ministry of Labour and Social Affairs of Ethiopia (MOLSA, 1997) and the number of rural school attendants decreased as the grade level increases (Bureau of Education of ANRS (BOE), 1997).

The average landholding per rural household was only 1.0 hectare (CSA, 1998). According to this very same source, the yield per hectare is one of the lowest and the largest proportion of rural farmland (30.29 %) was cultivated for *teff* production. Nowadays, in peasant agriculture the size of land per person is decreasing mainly due to ever increasing number of the rural population.

Though unemployment is not of serious concern, unlike the urban areas, a report by CSA (1992) revealed that nearly 25 percent of males and 28 percent of females, and about 26 percent of the total rural population were underemployed on the basis of the criteria "those who worked for less than the usual number of days".

More than half of the total cultivated land of the region was intensively and moderately cultivated, and the land covered by forest was only 1.5 percent of the overall area of the regional state (Bureau of Planning and Economic Development of ANRS (BOPED), 1997). It was estimated that about 1.1 billion tones of soil (58 % of the country's soil resource taken away by erosion annually) has been eroded by flood annually. According to BOPED reports population explosion, deforestation, and plateau-dominated topography of the region are the root causes of the problem

The rural population is involved in some non-farm activities mostly in slack times. Of the total rural households engaged in non-farm activities in the country about 11 percent of them were found in the region (MOLSA, 1997). The largest proportion was engaged predominantly in weaving (41%), pottery (13%) and spinning (9%). Weaving, which is usually practiced by Muslims, is a common non-farm activity in areas such as North Gondar and North Wollo (Relief and Rehabilitation Commission (RRC), 1994). Artisan activities, especially blacksmithing, pottery and tannery (which have long history in some parts of the region), are also practiced by rural households.

Recurrent drought, massive land degradation (as a result of population growth) and lack of basic infrastructures for intensive land use are some of the observable characteristics of most of the Amhara region (Mulat & Teferi, 1996; BOPED, 1997).

This study is divided in to five chapters. The first chapter introduces briefly the overall situation of traditional agriculture in Africa and the background of the study area. Objective of the study and the method of analysis are dealt in chapter two. The third chapter deals with the different reviewed studies on the role of non-farm activities and on some determinants of these activities. The logistic regression model results and the corresponding discussions are presented in the fourth chapter. The final chapter is about the conclusion and the possible recommendations of the study. Some of the logistic regression model diagnostic results are also appended.

OBJECTIVE AND METHODOLOGY

2.1 Objective

Studies show that a large number of households are involved in RNFAs¹ though the return that they are getting is not promising. Recent studies in Ethiopia revealed that farm households were diversifying their income sources as agricultural income is falling and becoming increasingly risky (MOLSA, 1997).

The non-farm sector in Ethiopia is not without hindrances. Besides the problem on agriculture, RNFAs are hampered, for instance, by absence/lack of physical and social infrastructures (lack of market and road accessibility, lack of availability of water, etc.), lack of credit facility/service and land, lack of skill and knowledge and target oriented policy directed to encouraging the sector (Mulat and Teferi, 1996; MOLSA, 1997).

However, little is known about the way these and other variables influence involvement in RNFAs in the country in general and in the Amhara region in particular.

Thus, this study attempts to identify the most important factors influencing non-farm employment in the rural Amhara region. The specific objectives are:

- To examine the effects of some variables that may influence involvement in RNFAs and identify the most important determinants by employing statistical methods,

¹Rural non-farm activities (RNFAs) in this study are defined as any income generating activity beyond agriculture performed by a person in rural areas. The words Non-agricultural, Non-farm and Off-farm are used interchangeably in this study.

- To describe the situation of involvement in RNFAs by categorising the predictors into different groups,
- To describe households involved in RNFAs by classifying them into different wealth groups,
- To provide information for the development of policies and studies regarding RNFAs.

2.2 Data Source

The data for this study are secondary data which are part and parcel of a national survey carried out, from the 3rd week of Oct. 1996 to the 1st week of Dec. 1996, by the Ministry of Labour and Social Affairs of Ethiopia entitled " Agricultural Wage Employment and Rural Non-Farm Employment Survey" financed jointly by UNDP and FAO. The survey covered rural areas of five regional states: Afar, Amhara, Oromiya, Southern Nations, Nationalities and Peoples and Tigray.

A two-stage stratified sampling design was applied for the survey - the first stage being selection of rural *kebeles* (the lowest administrative unit of Ethiopia's current administrative structure) and the second selection of households within the sampled *kebeles*. The *kebeles* were selected with probability proportional to measure of size (ppz) of households of *kebeles*. Thirty households, ten of them were for the non-farm study, were selected systematically from each of the selected *kebeles*.

The number of *kebeles* (from the list or frame of *kebeles* developed) to be selected from each region was first allocated to the zones into which the region has been sub-divided proportionally to their household estimates. As households involved in RNFAs are usually small in an

agrarian economy, a large area unit was adapted in the first stage of the selection to capture this situation.

Prior to the interview a fresh list of households was prepared in each of the selected *kebele*. The list was used to stratify rural households (into those involved in RNFAs, wage employment and agriculture only participants, for sample selection. Stratification at the household level was introduced, rather than using simple random sampling, mainly to reduce the loss of sampling efficiency due to the small proportion of non-agricultural population in an agrarian economy.

About 5,699 households were successfully interviewed at the national level. Of these 1,346 (nearly 30.3 percent of these were involved in RNFAs), which were selected from forty-five *kebeles* covered by the survey in the Amhara region, are made the focus of this study.

2.3 Variables of the Study

Separate data on handicraft, food and drink, and trade and small-scale industry activities were available in the data source. These can be used to fit separate multivariate analyses for identifying the important influencing factors of engagement in the three groups of rural non-agricultural activities. However, due to the small number of households involved in the activities and a large number of missing values in the data it is found meaningful to combine the three groups of RNFAs into one response variable so as to get an overall regional insight about the influencing factors of involvement in RNFAs in the Amhara region. The variables that are supposed to predict involvement in RNFAs are selected based on experiences from the revised studies and the available data on the subject.

2.3.1 Response Variable

The response variable, which is binary, provides data on number of rural households undertaking handicraft and related, food and local drink, and trade and small-scale industry activities and farm households that were not involved in any of these activities. The variable, denoted as RNFA, is represented as follows:

RNFA – Involvement in non-farm activities.

- 1**- Households with at least a member of the household involved in undertaking handicraft, food and local drink, trade and/or running small-scale industry activities only or along with agricultural activities
- 0** - Households involved in agriculture only

2.3.2 Predictor Variables

There is an unavoidable dilemma when some continuous covariates are modeled linearly in the logit (Hosmer and Lemeshow, 1989). It is recommended that dichotomizing a continuous variable is helpful whenever the change in the **log-odds** of the outcome variable as a result of a unit change in a particular predictor variable is not sound.

Most of the variables in this study are related to income and some are demographic. It is usually the case that data on income are skewed highly likely due to misreporting; age, for instance, is not free of preference besides the problem with the respondents in remembering exact birth dates especially in an illiterate rural society. Reports on the other variables, such as household size, are most likely dependent on the economic status of a household and the prevailing situation of the area during the survey period.

Due to these and other probable biases in the data and to make interpretations simpler and understandable the predictor variables of the study are made dich- or polytomous. Categorisation of the values of some of the predictors was based on their means or quartiles.

The regressors are grouped into three: demographic, endowment and other variables.

DEMOGRAPHIC VARIABLES:

RELIGION- Religion of Household Head

1 - Orthodox Christians

0 - Others

MAR_STAT- Marital status of Household Head

1 - Married

0 - Otherwise

P-ORIGION –Place of Origin of Household Head

1 - Born in ANRS

0 - Otherwise

HH-SIZE - Number of Household Usual Members

0 - ≤ 5 Members

1- ≥ 6 Members

SEX – Sex of Household Head

1 - Male

0 - Female

AGE- Age of Household Head in Complete Years

1- 18-30 Years

3- 41-50

2- 31-40 Years

4- >50

EDU_HEAD- Educational Attainment of Household Head

1 - Literate

0 - Illiterate

WEALTH VARIABLES:

FARM_SIZE- Size of Farm Land(in ha) per Household Member

1- 0-0.0625

3- 0.1876 –0.3738

2- 0.0626 – 0.1875

4- ≥ 0.3739

YIELD –Annual Agricultural Yield (in qt) of a Household

- | | |
|------------|-----------------|
| 1- 0 - 3 | 3- 6.01- 11 |
| 2- 3.01– 6 | 4- ≥ 11.01 |

NUM_LIVS – Total Livestock Number of a Household

- | | |
|----------|-----------|
| 1- 0 – 2 | 3- 7-11 |
| 2- 3 - 6 | 4- > 11 |

LIV_INC- Livestock Income(in Birr) of a Household

- | |
|--------------------|
| 0 - ≤ 39.6452 |
| 1- > 39.6452 |

ADD_INC – Amount of Additional Income(in Birr) of a Household

- | |
|-----------------|
| 0 0 - 68.5509 |
| 1 ≥ 68.551 |

LABOUR- Number of Daily Labourers hired for farm work

- | |
|------------|
| 0 - 0 –1 |
| 1 - 2 - 81 |

BORROW – Amount of Money Borrowed(in Birr) from Local Lenders

- | |
|------------|
| 0 - 0 - 48 |
| 1- > 48 |

FOOD_BAL –Food Balance(grain sale less grain purchase) in qt of a Household

- | | |
|-------------------|--------------|
| 1- -18.5- -2.2625 | 3- -0.79 – 0 |
| 2- -2.2624 - -0.8 | 4- > 0 |

LAND-OWN - Farm Land Ownership of a Household

- | |
|--------------|
| 1 - Owner |
| 0 - Landless |

DRAFT- Shortage of Draft Animals of a Household

- | |
|---------|
| 1 - Yes |
| 0 - No |

FERT - Fertiliser use of a household

- | |
|--------------|
| 1 - Users |
| 0 - Nonusers |

CUL_OTH – Whether a Household Rent in Farm Land

- | |
|--------|
| 1 -Yes |
| 0 - No |

ADD_SOUR- Presence of Additional Source of Income

1 - Yes

0 - No

PAID_WKR – Paid worker Hired for Herding, Farm Work or Household Work

1 -Yes

0 - No

LIV_HUSB – Engagement in Livestock Husbandry

1 - Engaged

0 - Not engaged

CROP_PROD – Engagement in Crop Production

1 - Engaged

0 - Not engaged

LAB_SHOR – Shortage of Labour for Farm work

1 -Yes

0 - No

OTHER VARIABLES:

MAIN_EMPS – Main Employment Status of Household Head

1 - Employer

2 - Own account Worker

3 - Employee

DIS_MAR - Distance of a Household(in hrs) from the Nearest Market

1- 0 - 1

3- 2.01- 3

2- 1.01- 2

4- > 3

DIS-ROAD-Distance of a Household(in hrs) from the Nearest All Weather Road

1- 0 - 1.3

3- 3.01- 8

2- 1.4 - 3

4- > 8

AGExFOOD_BAL- Interaction between **AGE** and **FOOD_BAL**

Households headed by older heads and with sufficient agricultural reserve are expected to be engaged in agriculture than in RNFAs.

YIELDxNUM_LIVS- Interaction between YIELD and NUM_LIVS

Households with large number of cattle and sustainable agricultural income are less likely involved in income diversification.

YIELDxHH_SIZE- Interaction between YIELD and HH_SIZE

Rural households who are food insufficient and are having more household size are expected to diversify their income to supplement subsistence livelihood.

DIS_MAR x DIS-ROAD –Interaction between DIS_MAR and DIS-ROAD

Roads and market places initiate non-farm activities to expand and new ones to come up. Moreover, they have been serving as good sources of income for some rural households. Under normal circumstances households in the vicinity of these infra-structures are highly likely in RNFAs than those far away.

FARM_SIZE x HH-SIZE – Interaction between FARM_SIZE and HH-SIZE

Usually the main activity of households in the rural areas is agriculture and hence the largest share of households' labour is expected to be allocated to farm activities. It can be hypothesised that farmers who are satisfied with their agricultural output are less likely interested in income diversification. This actually may depend on the educational level and/or training and other personal attributes.

2.4 Methodology

2.4.1 Comparative review of some Qualitative Response Models (QRM)

Fitting a multiple linear regression model of a dichotomous response variable on a set of continuous and/or dichotomous regressors encounters three basic difficulties: violations of the

fundamental least squares assumptions- homoscedasticity and normality of errors, and the predicted values that the regression analysis yield are not guaranteed to lie in the range of 0 to 1 (Montgomery and Peck, 1992; Sincich, 1993).

A number of alternative statistical models, with different merits and shortcomings, are developed for analysing qualitative response variables. Logistic Regression (Sincich, 1993; Montgomery and Peck, 1992; Agresti, 1992 and others), Discriminant Function (Rubin, 1984; Hosmer and Lemeshow, 1989), and Pobit model (Desta, 1990) are some of them.

Rubin (1984) indicated that discriminant function analysis, though it makes computations of the estimators easier, is not preferred for fitting a qualitative response variable due to sensitivity of the discriminant function estimators to the assumption of normality and the overestimation of the association between the outcome variable and the regressor under question that follows, especially when the variable is dichotomous. Moreover, more parameters need to be estimated in discriminant function analysis than in logistic regression model.

Linear probability (linear regression) with proportion of success as the outcome variable could be used as a model showing the degree of effect of each regressor on the probability of the response variable falling into the categories.

The ceiling/floor problem of linear probability model can be overcome by replacing odds-ratio in place of proportions. But this too has its own shortcoming in that we cannot logically state the effect of a regressor on the odds is linear as factors affect the odds are multiplicative.

The other option for analysing qualitative response variables is the probit (abbreviation for probability unit) model. The model, which is the inverse of the Gaussian distribution, is de-

rived from the definition of a cumulative density function of a random variable, say X ($P(X \leq x) = F(X) = \int_{-\infty}^x f(u)du$, where $F(\cdot)$ is the distribution function and $f(\cdot)$ probability density function of X).

Collet (1991) described that this model has similar type of limitations and advantages with logistic regression but differs from the latter in the nature of transformation employed.

Graphical analyses showed that it is usually difficult to discriminate between the probit and logit functions on the basis of goodness-of-fit especially over the interval $0.1 \leq p \leq 0.9$, where p is probability of success (McCullah and Nelder, 1989).

Amemiya (1981), cited in Desta, has shown that the probit, logit and linear probability (lp) models are related as follows: $1.6b_{probit} \approx b_{logit}$ (in general); $b_{lp} \approx 0.4b_{probit}$ (for model parameter estimates other than the constant term) and $b_{lp} \approx 0.4b_{probit} + 0.5$ (for model constant term estimate); $b_{lp} \approx 0.25b_{logit}$ (for model parameter estimates other than the constant term) and $b_{lp} \approx 0.25b_{logit} + 0.25$ (for model constant term estimate), where the b 's are parameter estimates of the models correspondingly mentioned.

However, in some cases, logistic regression is preferred to the probit due to its link to other models, such as log-linear models, and its simpler interpretability as the logarithm of the odds ratio and its eminence for retrospectively collected data analysis (McCullah and Nelder, 1989). Moreover, unlike the polytomous logistic regression, the multinomial (i.e., more than two levels for the dependent variable) extension probit model is logically possible but impractical.

The problem of using odds, as they are, as values of a dependent variable in modeling a qualitative outcome variable can be alleviated by taking the logarithm of the odds as a response variable in the regression analysis (McCullah and Nelder, 1989). Such types of models, called logit (logarithm of the odds) help to do away with the ceiling/floor constraint and enables the parameter estimates of the logistic model to state the linear effect of each independent variable on the logit of the odds. Logistic regression (logit) is, then, useful in order to model data that do not suit the linear probability regression.

Hosmer and Lemeshow (1989) pointed out that logistic regression model (and hence the logit) is a remarkably flexible model and it is unlikely that any alternative model provides a better fit for quantal response variables if the data is free from outliers and other inconveniences. The interpretability of the differences on the logit scale whether the data are sampled prospectively or retrospectively makes the model distinguishable from the other qualitative response models (McCullah and Nelder, 1989).

2.4.2 CHOICE OF MODEL

As it has been indicated in the brief comparative review section above, there are different statistical methods for assessing the effect of one or more continuous/discrete regressors on a dichotomous outcome variable. Some of these are the discriminant function analysis, the probit model and the logistic (logit) regression. Of these the logistic regression model, which is concerned with the probability of occurrence of an event, is widely applied mostly in social sciences and epidemiology, it is, nowadays, being considered as a standard method of data analysis in epidemiological studies (Lemeshow and Hosmer, 1982). Not only the availability of well established estimation methods but also its ease of interpretation as well as its close relationship to log-linear analysis of contingency tables and linear discriminant function analysis make the logistic model more popular than other related models.

represents the conditional probability that an outcome is present given the $\mathbf{X}$ -set. P_i is explicitly expressed as

$$p_i = P(X) = P[y_i = 1 / \mathbf{X}] = \frac{\text{Exp}(\mathbf{X}\beta)}{1 + \text{Exp}(\mathbf{X}\beta)} = \frac{1}{1 + \text{Exp}(-\mathbf{X}\beta)} \quad (\text{Multiple Logistic Regression Model})$$

Model)

Logarithmic transformation yields the Logit Regression Model,

$$\ln\left\{\frac{p(\mathbf{X})}{1-p(\mathbf{X})}\right\} = \mathbf{X}\beta = \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}, \text{ where } \beta \text{ is a } (p+1) \times 1 \text{ vector of unknown coefficients.}$$

coefficients.

Since y_i 's ($i = 1, 2, \dots, n$) are independently distributed as binomial with parameter p_i , $\text{bin}(1, p_i)$, the likelihood function of Y is expressed as

$$\ell(\beta, Y) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} = \prod_{i=1}^n \left(\frac{1}{1 + \text{Exp}(-\mathbf{X}\beta)} \right)^{y_i} \left(\frac{\text{Exp}(-\mathbf{X}\beta)}{1 + \text{Exp}(-\mathbf{X}\beta)} \right)^{1-y_i}$$

Our objective here is to obtain estimates of the vector $\beta^1 = (\beta_0, \beta_1, \dots, \beta_p)$ applying the ML estimation technique.

The method of ML estimation proceeds by considering the natural logarithm of the likelihood

$$\text{function}(\ell(\beta; Y)), \quad LL = \ln \ell(\beta; Y) = \sum_{i=1}^n \{Y_i \mathbf{X}_i \beta - \ln(1 + \text{Exp}(\mathbf{X}_i \beta))\}$$

and taking the partial derivatives of LL with respect to the $p+1$ parameters and equating each to zero. This results in $p+1$ likelihood equations with $p+1$ unknowns. As these likelihood equations are nonlinear in the parameters, the estimation of the coefficients proceeds with numerical iterative techniques.

The second partial derivatives

$$\left(\frac{\partial^2 LL}{\partial^2 \beta_j}\right) = -\sum_{i=1}^n x_{ij}^2 p_i (1 - p_i) \text{ and } \frac{\partial^2 LL}{\partial \beta_j \partial \beta_u} = -\sum_{i=1}^n x_{ij} x_{iu} p_i (1 - p_i)$$

which together called a $(p+1) \times (p+1)$ Information Matrix denoted as $\mathbf{I}(\beta)$, gives the variance-covariance matrix $\hat{\mathbf{I}}^{-1}(b)$ of the estimates (Hosmer and Lemeshow, 1989). The iterative ML estimation method (known as the Newton-Raphson method) expresses b at the $(t+1)^{\text{st}}$ iteration as $b(t+1) = b(t) + (\mathbf{X}'\hat{\mathbf{V}}(t)\mathbf{X})^{-1} \mathbf{X}(t)r(t)$, $t = 1, 2, \dots$

where $\hat{\mathbf{V}} = \text{diag}(\hat{P}_i(1 - \hat{P}_i)) = \text{Cov}(\mathbf{Y})$, $\mathbf{b}$ is the resulting ML estimator of β and $\mathbf{r} = \mathbf{Y} - \hat{\mathbf{P}}$ (Landweh et al., 1984).

All the parameters $\beta_0, \beta_1, \dots, \beta_p$ can be estimated employing the appropriate sub-routine in SPSS and then estimates of $P(y=1/\mathbf{X})$ for each subject under study can be computed accordingly.

As far as interpretation of coefficients is concerned, each logistic regression coefficient has a straightforward interpretation. The estimated coefficients, b_j 's, represent the change in the log-odds for every unit (c units depending on the particular variable) change in x_j holding all other independent variables constant, and $\exp(b_j) = \hat{P}_i / (1 - \hat{P}_i)$, $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, p$, which is the odds of being in the category of interest for the i^{th} subject, represents the multiplicative factor by which the odds change for every unit (c units) change in x_j controlling for the other predictors. The coefficient of a two category indicator regressor, for instance, is interpreted as the difference between the log-odds when a member of success (coded as 1) category and failure (coded as 0) category (or equivalently, the amount of change in log-odds as a result of the change in the predictor variable from 0 to 1) considering the other independent variables in the model remain unchanged. Similar interpretations can also be made for a

regressor variable having $k(k>2)$ levels by making one of the k levels as a referent category (Hosmer and Lemeshow, 1989).

2.4.4 Variable selection and Goodness- of - Fit Assessment of the Model

Once the multiple logistic regression model is fitted assessment of the fit is the next task to go through assuming reliable estimates are chosen in the estimation stage.

The assessment of goodness-of-fit of a logistic regression model is concerned with checking the functional form of the model through testing the influential nature of each candidate predictor to the regression. But before that it is usually advisable to conduct univariate analyses (both graphically, for continuous covariates, and significance tests) to see the relationship between each competing regressor and the outcome variable.

Goodness-of-fit of the model can be assessed by seeing how well the model classifies the observed data (in the Classification Table) or by examining how “likely” the sample results actually are, given the estimates of model parameters (SPSS, 1994; Hosmer and Lemeshow, 1989).

In the former case, close examination of the confusion matrix, by modifying the threshold value whenever necessary, will help to judge the overall performance of the model. The fit is considered to be good if the overall correct classification rate exceeds 50 percent. The choice of a suitable threshold value is made either by identifying the value that minimises the overall proportion of misclassification or by compromising between the minimisation of the two misclassification probabilities, namely the probability of declaring an individual to be in group 0 when it should be in group 1 and vice versa (Collet, 1991).

While in the latter approach (above), we can measure how well the multiple logistic regression model fits via the likelihood ratio statistic (-2LL) or the goodness-of-fit statistics

$(X^2 = \sum_{i=1}^n \frac{(Y_i - \hat{P}_i)^2}{\hat{P}_i(1 - \hat{P}_i)})$. The model is a good fit if the difference between the observed probabilities and predicted values are close to zero, i.e., X^2 or -2LL is very small compared with $\chi^2_{(n-p)}(\alpha)$, where p is the number of predictor variables (including the constant term, β_0) in the model. Complete assurance of goodness-of-fit is possible only after thorough residual analysis is made.

Testing the null hypothesis that the p coefficients for the covariates in the model are equal to zero, proceeds in the likelihood ratio case as follows:

$H_0: \beta = 0$ (all of the p covariates are not important in explaining the outcome variable).

$H_1: \beta_i \neq 0$ (at least one of the covariates is important in explaining the outcome variable).

Let L_0 and L_1 be the maximized log-likelihoods under H_0 and H_1 , respectively.

The computed values $-2LL = -2(L_0 - L_1)$ which is distributed (under H_0) as χ^2_p serves as a test statistic. If the value -2LL, which is the model chi-square value in this case, is larger than $\chi^2_p(\alpha)$ then at least one of the p -variables included in the model are important in explaining the response variable.

The next task in the assessment of importance of the variables, if the null hypothesis above is not rejected, is investigating which of the p -covariates best explains the response variable, and for this we carry out univariate analyses.

The test for the significance of each predictor (i.e., $H_0: \beta_i = 0$ versus $H_1: \beta_i \neq 0$ $i = 1, 2, \dots, p$)

can be undertaken using the Wald statistic
$$\left(\frac{\sum_{i=1}^n x_i (y_i - \bar{y})}{\sqrt{\bar{y}(1-\bar{y}) \sum_{i=1}^n (x_i - \bar{x})^2}} \right).$$

However, Wald's test (though it requires fitting the model only under the alternative hypothesis in contrast to the likelihood ratio test which requires fitting the model twice) has undesirable properties: the test statistic decreases to zero as the distance between b_i and β_i increases, and the power of the test, based on its large-sample distribution decreases to the significance level (α) for alternatives sufficiently far from the null value (Hauck and Donner, 1977). As far as efficiency in testing the coefficients of a regression is zero is concerned the score statistic $(\hat{\mathbf{R}} \hat{\mathbf{I}}^{-1}(\mathbf{b}) \hat{\mathbf{R}} \sim \chi^2_p$, where $\hat{\mathbf{R}}$ is a vector whose i^{th} element is

$\frac{1}{\sqrt{n}} \frac{\partial \ell(\beta, y)}{\partial \beta_i} \Big|_{\beta_i = b_i}, i=0,1,2 \dots p$) (Rao, 1973) is not more efficient than the Wald statistic. But

the three test- statistics are all equal in large samples, when the null hypothesis is true (SPSS, 1994).

Variable selection is one part of multiple logistic regression model building process where subject matter knowledge and parsimony are given due consideration.

As is the case of multiple linear regression, the logistic regression technique has as such no 'best' variable selection method. However, careful assessment of the data through the different methods may yield an appropriate model that fits the data well. In this case too it is helpful to perform univariate analyses regarding the relationship of each regressor with the response variable before the selection of the final model (Hosmer and Lemeshow, 1989; Agresti, 1996). A two sample t-test (due to its close relationship with linear discriminant function) or a univariate logistic regression and a likelihood ratio chi-square test (for a $2 \times k$ contingency ta-

ble) can be employed to see the importance of a continuous variable and a k- level variable to the outcome variable. The importance of a continuous covariate to the logistic regression can further be confirmed by a scatter plot on the logit scale as Copas (1983), cited in Hosmer and Lemeshow (1989), suggested.

According to Hosmer and Lemeshow (1989), all variables whose univariate tests have a p-value smaller than 0.25 and supported by the plot and those with known demographic importance can be taken as potential candidates for the multivariate logistic regression model. Any regressor with zero cell frequency when cross-classified with the outcome variable should be modified or discarded before running the logistic program.

The **forward/backward stepwise** and **Enter** methods are the different variable selection methods available to select important predictors. Comparison of the outputs of the forward/backward (-conditional, -likelihood ratio, -Wald) stepwise methods is helpful as to which variable to select so that the model we are looking for is meaningfully fit and parsimonious.

A decision on the importance of each variable included in the particular variable selection method is made based on examination of the Wald statistic for the variable, and comparing each estimated coefficient of the particular variable on the multivariate model with the univariate estimate of the model containing only that predictor. Accordingly, adjustments (deletion, refitting, and/or verifying) can be performed, on the basis of thorough investigation on the estimates of the parameters, and the process continues till all the important predictors are identified. Having selected all the variables that are statistically significant, possible interaction terms between the regressors remaining in the model can be considered and tested.

'Best' logistic regression fit with individually and jointly important regressors can be obtained after thorough objective and subjective scrutiny of univariate results and comparing the outputs of the forward/backward stepwise and other variable selection methods.

2.4.5 MODEL ADEQUACY

Evaluating the adequacy of a statistical model is an indispensable step of a multiple regression problem. The adequacy of a fitted regression model may fail due to the presence of one or more extreme observations, influential cases and/or complete (near) linear dependencies among the regressors.

That means examination of residual measures before concluding that the model fits is crucial. This section aims at establishing the methodology for investigating the validity of the logistic regression model: detection of outlier(s), influential observation(s) and examination of inter-correlation between the different predictors in the model.

2.4.5.1 Detection and Treatment of Outliers

An outlier, in logistic regression setup, shall mean an observation with a success ($y=1$) with probability of success ($P(y=1/X)$) close to zero or an observation with a failure ($y=0$) with probability of success ($P(y=1/X)$) close to one (Jennings, 1986; Copas, 1988). This is due to the inclusion of extreme observations in one or more of the explanatory variables.

According to Copas (1988), outliers in binary regression can only occur towards the edge of the design space on graphical displays except when points that lie near the boundary of the x-space are either all 0s or all 1s.

The unstandardised residuals, $Y_i - \hat{P}_i$, the standardized residuals, $\frac{(Y_i - \hat{P}_i)}{\sqrt{\hat{P}_i(1 - \hat{P}_i)}}$, the studentised residuals (the change in the model deviance if the case is deleted), Deviance

$\pm 2\sqrt{\ell n\left\{\left(\frac{y_i}{\hat{p}_i}\right)^{y_i}\left(\frac{(1-y_i)}{(1-\hat{p}_i)}\right)^{1-y_i}\right\}}$, and the Half-Normal Plot of standardized residuals are some of the statistical tools to see the presence and extent of influence of outliers.

Any observation that appears to be quite unusual on any of the case-wise plots of residuals or deviance and Half-Normal Probability plot of standardized residuals warrants examination. The corresponding decision whether to discard a suspected observation is based on the result of further investigation on the particular observation.

2.4.5.2 Influence Diagnostics

One or more subjects in the independent variables set may put influence on the regression model parameter estimates and also on test-statistics and confidence intervals which are functions of the estimates. The influence will be exacerbated if the observation is an outlier.

The matrix $H = \hat{V}^{1/2}X(X'\hat{V}X)^{-1}X'\hat{V}^{1/2}$, where $\hat{V}$ is an $n \times n$ diagonal matrix whose $(i,i)^{th}$ entry is $\hat{P}_i(1 - \hat{P}_i)$, usually called the “hat” matrix plays the central role in influence analysis (McCullah and Nelder, 1989; Agresti, 1996). Close examination of each of the diagonal elements of H , $\hat{h}_{ii} = (1, X_i)(X'\hat{V}X)^{-1}(1, X_i)'\hat{P}_i(1 - \hat{p}_i)$, helps to identify leverage value(s) in the X -set. Hoaglin and Welsh (1978) cited in McCullah and Nelder (1989) suggested that the i^{th} subject is a leverage value if $\hat{h}_{ii}$ exceeds $2p/n$, where p is the number of independent variables in the model. The relationship between leverage and predicted probabilities will also enable us to view leverages from different perspectives. According to Hosmer and Lemeshow

(1989), a particular value of leverage is considered to be influential if the estimated probability of success lies between 0.1 and 0.9; otherwise, that is, if the estimated probability is small (<0.1) or large (>0.9) then the leverage value may not have impact.

On the other hand, the influential nature of an observation can be assessed by comparing the results of a fit with and without the observation. The displacement of confidence intervals of regression parameters, *Dfbeta* (the change in the logistic regression parameter estimate when

an observation is deleted), *Cook's Distance*, $\frac{(Y_i - \hat{P}_i)^2}{\hat{P}_i(1 - \hat{P}_i)} \left(\frac{\hat{h}_{ii}}{(1 - \hat{h}_{ii})^2} \right)$, and the change in the

Pearson Statistic, $X^2 = \sum_{i=1}^n \left(\frac{(Y_i - \hat{P}_i)^2}{\hat{P}_i(1 - \hat{P}_i)} \right)$, or *goodness-of-fit statistic* ($-2LL$) when the obser-

vation is deleted can indicate whether an observation has a substantial impact.

If investigations of these changes revealed that one or more of the observations are so, implying the model with these observations is not adequate, then the logistic regression analysis proceeds with a reduced sample size.

2.4.5.3 Multicollinearity

Multicollinearity is near (complete)-linear dependence among regressors. If one of the predictors, say X_k , is expressible as a linear combination of the others, $X_1, X_2, \dots, X_{k-1}, X_{k+1}, \dots, X_p$ then there would result in singular $X'X$ (nonzero off-diagonal elements) and this in turn means that there is either no information or little information to estimate the regression coefficient of X_k given that the other variables are included in the model-leading to omission of X_k from the model without affecting the X -set (McCullah and Nelder, 1989; Montgomery and Peck, 1992).

Multicollinearity can occur due to poor sampling technique, incorrect specification or overdefining a model. Whatever the cause may be it is advisable to do multicollinearity diagnostic before proceeding with variable selection, inference or other regression analysis.

Near zero $\det(X'X)$ implies one or more large diagonal elements of $(X'X)^{-1}$, and this indicates which variables are included in the multicollinearity. The problem can be detected through a number of methods. Inspection of the off-diagonal elements of $X'X$, variance Inflation Factors (**VIFs**), i.e., diagonal elements of $X'X$, square root of **VIFs**, and condition indices (number) are the widely applied ones.

VIF for each predictor shows the combined effect of the dependencies among the regressors on the variance of the variable. Though it is not a rule, **VIFs** exceeding 5 or 10 indicates that the associated regression coefficients are poorly estimated. Square root of **VIF** is helpful in indicating how much the confidence interval of a regression coefficient is displaced due to multicollinearity. The condition indices are valuable in pointing the variables that are included in multicollinearity relationship. A condition number (index) between 100 & 1000 and larger than 1000 (in that order) are indications of moderate and severe multicollinearity.

Eliminating a regressor that is linearly dependent on other predictors in the model, with careful analysis of its consequences, helps to circumvent multicollinearity.

Chapter III

LITERATURE REVIEW

3.1 OVERVIEW OF STUDIES ON RNFAs

3.1.1 The Role and Trend of RNFA

Studies in Africa show that rural non-farm activities (RNFA) have been a central part of peasant economy and practice. About 28.0, 22.3, and 19.0 percent of the rural labour force in Kenya (1970*), Zambia (1980*), and Sierra Leone (1976*) were engaged in non-farm activities (ILO et al., 1988).

These activities are important means of income diversification especially when agriculture is in crisis. However, attention was not given in development research until very recently (Møller, 1998).

RNFAs are not uncommon in Ethiopia. According to MOLSA (1997), about 24 percent of rural households were engaged in weaving (of these 41 percent of them are in the Amhara region), 21 percent in basketry and 11 percent in pottery. However, most of these activities are carried out during some parts of the year and the income obtained from these activities is very small (Mulat and Teferi, 1996).

Mulat and Teferi revealed that an average annual income of Birr 480.65 was derived from traditional handicraft activities, and an average annual income of about Birr 980.92, 351.57, and 597.82 was obtained from the sale of local brews (*araki* and *tella*), trade and wage

* Shows the years when the data were collected from the respective countries.

employment activities respectively in North Shoa.

Tegegne (1996) documented that about 70.7 percent of the sample households in Kachabira and Damotgale *weredas* (Southern Nations, Nationalities and Peoples' State) obtained non-agricultural income. An annual income of Birr 633.95 per household was obtained in areas with predominantly trade and handicraft activities.

The findings of the above studies indicate that different RNFAs are being practiced in different parts of the country.

Møller (1998), based on a survey in two Nigerien villages, found that more than 92 percent of the households managed to get income from different non-farm activities. According to his findings, households who strongly integrated farming with RNFAs were found the most economically successful ones in overcoming the falling agriculture in the villages. It was also indicated that monetization of rural life, changes in family structures, introduction of new agricultural techniques, and low development in urban industry were the processes which have led to the adoption of RNFAs in peasant livelihood strategies of the villages.

Berkvens (1997) documented the growth of de-agrarianisation¹ its importance and implications in sub-Saharan Africa. He mentioned that the booming of RNFAs in southern Africa is mainly due to environmental degradation and declining of overall economic performance. According to his finding unemployment, Structural Adjustment Programmes (SAP) and stagnation of the formal sector were some of the pushing factors to RNFAs. It was further indicated that rural households are nowadays viewing RNFAs as dependable income source and as the best substitute/subsidy for the risky traditional agriculture.

¹ "... a process of economic activity reorientation, occupational adjustment and spatial realignment of human settlement a way from agrarian patterns" (Berkvens, 1997).

Berkvens pointed out that about 85 percent of the rural households of Mutoko Communal Area of Zimbabwe had some kind of non-agricultural cash income, which accounted, on the average, for 55 percent of total household's cash income. It was found out that the total share of RNFA income and the aggregate number of income sources per household has grown by 5% and from 2.15 to 2.81, respectively between 1985 and 1996 in the area. On the other hand, Gadidzanwa (1997) identified that Zimbabwean rural micro- and small-scale enterprises have been increasing at twice the rate of their urban counterparts since 1991.

Berkvens indicated that most of the activities in the area, like other RNFAs in Africa, have some characteristic in common-being one-man activity, located at homestead of farming household, operated by illiterates and/or less educated framers.

Secondary sources showed that the proportion of RNFAs income in total cash income per household rose from 25 percent in 1969 to 33 percent in 1977, 38 percent in 1980, and 47 percent in 1983 in rural Tanzania, and it has been known that about 87 percent of all rural households derive some cash income from RNFAs (Bagachwa, 1997). It was also found out that about 13% of the country's economy value-added in 1991 was from the rural informal sector. The overall situation of the sector is similar to that of Zimbabwe and Niger.

Jambiya (1998) revealed that the number of RNFAs participants reached over a third of Usambara and Lushoto population in the periods 1980-96. Moreover, it was found out that about 24.4 percent of RNFA actors generated a substantial proportion of their household income from these activities.

Mung'ong'o (1998) explained that the evolution of RNFAs in Njombe district (Tanzania) was promising. He expressed that non-agricultural activities, which are predominantly performed during slack period of agricultural seasons, were widely practiced in the area. It was

stated that the number of RNFA participants in Njombe district had increased by about 30 percent in the years 1980-95. A tenfold increment in the number of RNFA actors in handicrafts such as knitting of cloth and weaving of mats and ornaments and a 33.3 percent growth in brewing and selling local beer had been observed in the district during this period.

Ibrahim (1997) mentioned that remittance, migrant works, and engagement in different handicraft and trade activities were common in Sudan due to the influence of SAPs, shrinkage of the state employing organisations and drought.

As Maegher and Mustapha(1997) explained over one-half of the total cash income of small-scale farmers in Hausa, Nigerian Province, was derived from RNFAs. Moreover, it was pointed out that the overall share of RNFAs income from the total cash income of farm households had shown increment in the region in the years 1974/5, 1989/90 and 1992/3. On the other hand, it was revealed that wealthy, middle and poor echelons of rural households, in the vicinity of Gusau and Sokoto towns of the country, enjoyed non-farm incomes amounting to between 48 and 60, about 14 and 60, and 33 and 58 percent of their incomes respectively (Iliya and Swindel, 1997).

About 71 percent (1.5 persons per household) of the economically active force of the inhabitants of Abeka, Ghana, were in the rural informal sector (Van Dijk, 1997). According to Foeken (1997) the situation of RNFAs in some parts of Kenya was similar to the case of Abeka. About 1.4 members per wealthy and middle class households in the Coast Province of the nation were engaged in some type of off-farm activities and more than 60 percent of their income was derived from RNFAs.

3.1.2 DETERMINANTS OF RNFA

The type and significance of RNFA's determinants are not explicitly and exhaustively defined in the literature. This may be probably due to low attention given to the sector in development research. In many situations, researchers viewed that factors influencing agriculture, such as natural hazards, market fluctuations, social uncertainties and state actions and war, in one way or the other, put influence on RNFA. Some of the literature, which are believed to be better elaborated and related to the issue, are reviewed below.

Møller (1998) described the influencing factors of RNFA as seasonal, personal and general. Seasonal factors mostly encompass climatic fluctuations between years; personal factors refer to household's socio-economic status, and general relates to natural hazards (drought, floods, etc.), political and economic state actions. Location is one important determinant of non-agricultural activities. It was observed that the longer the distance of RNFA, pottery for instance, to urban centers (market places), the less the urban character of the product and the smaller the technological input in them. The closer to urbanised areas, the stiffer the competition from lower-priced, manufacturing substitutes from the urban informal sector and/or the formal enterprises. RNFAs were also mentioned as dependent on time.

According to Møller the degree of income diversification is markedly positively correlated with household heads' age. It was explained that, households headed by older heads are most likely interested in RNFAs than younger ones.

Non-farm earnings and landholdings were found to be inversely related in the Indian semi-arid areas, northern Nigeria, and Sierra Leone (Reardon et al., 1992). But studies in other parts of India revealed that this relationship is U-shaped; it was also found out in Burkina Faso, for instance, that greater landholding means greater income diversification in deficit ar-

areas whereas more land means less income diversification in surplus producing areas (Reardon et al., 1992). The same source revealed that demographic variables such as household size, age of head were found to have no influence on income diversification in Sahelian, Guinean and Sudanian tropical zones of Burkina Faso.

Comprehensive model based studies were conducted on RNFA in Ethiopia in recent years. These studies identified important determinants of non-farm employment in the country.

Mulat and Teferi classified the influencing factors of RNFA as personal attributes, farm income, land endowment, and food balance (grain sale less grain purchase). They fitted a multiple linear regression model of the ratio of non-farm income to total cash income of households on the following predictors: age of the household head (Age), Sex of the head (Sex), Educational Level of the head (Education), Family Size, Income from the sale of crop (Crop Revenue), Income from sale of Livestock and related (Livestock Revenue), Total agricultural Yield (Yield), Farm Size (Holding Size), and Food Balance at the household level.

The summary result of the study, based on the whole sample, showed that Yield, Holding Size, Food Balance, Livestock Revenue and Age of the head were identified as the only significant predictors of the share of RNFA income from the total household's cash income in North Shoa.

The negative sign of the coefficient of the variable Age in the fitted model, as opposed to Møller's (1998) finding, indicates that the share of non-farm income was higher for younger farmers than older ones. On the other hand, the negative signs of the coefficients of Yield, Holding Size, Food Balance and Livestock Revenue in the estimated regression model indicates that the share of non-farm income from the total household income was lower for households who were satisfied with agriculture.

MOLSA, based on a national survey on five administrative regions, analysed the influence of some endowment, demographic and other relevant variables on probability of participation in handicraft, food and drink, and trade activities. Separate multiple logistic regression models were fitted for each of the three dichotomous outcome variables.

The findings of the analysis revealed that trade, food and drink, and handicraft activities were negatively and significantly influenced by farm related variables such as, number of cattle, farm size per household member, and fertilizer use. This result agrees with Mulat and Teferi's finding reviewed above. Income from different agricultural activities was found to have a negative and significant influence on participation in handicraft activities and food and drink services but a positive and significant effect on engagement in trade employment.

With regard to demographic determinants, it was found out that female-headed households were active participants in the three types of RNFAs. Educational attainment of heads positively influence participation in food and drink service and trade RNFAs but negatively on handicraft employment.

Infrastructural determinants (distance to the nearest market place and all-weather road) were found significant influencing factors of the three non-farm activity types. The separate logistic regression results showed that the closer farm households to market places and roads the higher the probability of involvement in handicraft, food and drink and trade activities.

Income from RNFAs (mainly from trade and handicraft) was found to be positively correlated with cultivated land size, cost of fertiliser and crop sale revenue for households engaged in non-farm employment only as opposed to the combined sample of farm and non-farm households (Tegegne, 1996). This explains that non-farm income has enabled farmers to get more

cultivated land through share cropping and renting and also served as a supplement for purchasing farm inputs thereby improving their living standards. It was also indicated that income from handicraft and trade activities positively determined fertiliser use and crop sale revenue of non-farm actors.

A multiple linear regression model, similar to the one used by Mulat and Teferi on North Shoa was fitted by Tegegne. The regression result revealed that the demographic variables- age of the household head and family size have no influence on non-farm income, whereas educational level of heads of households was found to be significant and positively signed with non-farm income. The statistical analysis of the empirical data further disclosed that non-farm income was negatively and significantly determined by the proportion of farm income (crop sale plus livestock revenue) in total income of a household. These findings explain that households with strong base in the farm sector do not tend to engage in income diversification. However, the more farm households are skilled/educated, the greater the income from RNFAs.

3.2 Comments on Studies on RNFAs

Regression techniques are valuable tools for modeling relationships between a dependent variable and a series of predictors. Nowadays, model-based analyses are being increasingly used in various disciplines. The presence of well established model building strategies, self-dictating statistical packages, and access to high-speed computers are paving the way for the application of regression models.

Whatever methods of fitting we apply, a regression model has to be tested for adequacy and diagnosed. This includes checking the functional form between the outcome variable and the different regressors and model parsimony. Some summary statistics such as *adjusted R^2*

(a value that penalises the analyst whenever an unimportant predictor is included into a linear regression model) could be used to see the explanatory power of a particular model. Moreover, the standard errors of the regression estimators are also informative of the importance of regressors. Above all, residuals of a model and hence the different plots (of residuals versus case numbers) are essential tools to see the presence of extreme values and/or unduly influencing observations that in one way or the other make the model inadequate.

On the other hand, once the main effects of a model are found statistically significant and as long as meanings can be attached and significantly explain the dependent variable inclusion of interaction terms into the model should not be overlooked.

Much of the reviewed literature on RNFAs concentrated on quantifying the share of non-farm income in total income; identifying the factors (not in model terms) deriving households to diversify outside of agriculture; and examining equity and food security implications (Reardon et al., 1994).

The central tasks of regression analyses, goodness-of-fit tests and model diagnostic checks, are not treated in all of the reviewed model based studies on RNFAs.

Small *adjusted R^2* values were observed in most of the linear regression models reviewed. This may be probably due to insufficient number of explanatory variables and/or incorrect specification of the model. The presence/influence of high leverage values and outliers, which may be inevitable in model building processes, have not been dealt with in the multiple linear or multiple logistic regression models analyses of the reviewed studies.

The standard errors of regression coefficients, that appear in all of the models, which are indicators of imprecisely estimated parameters and sometimes overfitting (Collet, 1991), are used

for testing the significance of parameters only. Some of the reviewed models considered a regressor for their intended purposes whatever the test result may be.

As indicated previously, checking possible influences of observations on inferences, such as confidence intervals for model parameters, deserves attention while modeling relationships between variables. However, attempts were not observed in this regard in all of the model-based studies on determinants of RNFAs.

Some interaction terms could be hypothesised to be determinants of non-farm income. These terms might be interactions between household size and farm size, market distance and distance from all weather roads, and educational attainment of the head and farm income. Analyses on the effects of interaction terms were not included in the revised studies.

A model that best explains the probability of involvement in RNFAs will be fitted paying attention to the limitations indicated above.

Chapter IV

Results and Discussion

The data on involvement in RNFAs was checked for possible errors. While editing the data for analysis a number of problems were encountered. The presence of structured data, that is regressors having values almost identical to the outcome variable was one data problem that altered the maximum likelihood iteration procedure, and unusual figures given the overall demographic and economic situation of the area from where the data is collected are a few to list.

Model residuals were found unstable. That is, we encounter residuals that are numerous and higher in magnitude every time we run the logistic program even after the suspected outliers are removed. The considerable difference in the size of the two groups of the dependent variable and hence the effect on the estimated model was the other problem that took much time to arrive at a reasonable classification table.

Farm households with one or more missing values and those with unrealistic figures were excluded from both the univariate and multivariate analyses. This together with the discarded outliers reduced the size of the sample, for the multivariate analysis, from 1346 (408 involved in RNFAs and 938 engaged in agriculture only) to 1310 (375 involved in RNFAs and 935 engaged in agriculture only).

4.1 Univariate and Multivariate Results

4.1.1 Univariate Results

A preliminary assessment on the importance of each candidate predictor in explaining the probability of involvement in non-farm activities was made.

The univariate assessment is helpful in reducing the workload in selecting important predictors for the multivariate logistic regression model. It is also informative of the association between a predictor and the response variable.

The assessment emphasised not only on the significance of a particular regressor but also on investigating the magnitudes of the standard errors of model parameter estimates, for possible unreasonable inflation, and on the meaningful interpretation of the odds ratios under the existing situation of the survey areas. Attention was primarily paid to the assessment of the magnitudes of cell frequencies of each cross-classification of the dependent variable with the independent variables.

The list of demographic, endowment and related, and other variables that are found associated with involvement in RNFAs, at the univariate level, are displayed in Table1 below.

Personal attributes: age, sex, religion, and marital status of household head significantly determine involvement in non-agricultural activities. The univariate logistic regression results indicate that households who are headed by females and single, widowed and/or divorced farmers are found to be active participants in RNFAs. Muslims and/or other religious households are more likely involved in income diversification than Orthodox Christians. On the other hand, participation is found to be higher in households administered by relatively younger heads than older people when both are compared with households headed by young heads (age 18-30).

As regards endowment and related variables agricultural yield, number of livestock, rent-in farmland and labour shortage for farm activities are found to be influencing factors of involvement in non-agricultural activities.

Households, with higher agricultural yield levels are less likely to involve in income diversification as they might devote all the available household labour on farm activities. The preliminary assessment further revealed that farmers who hired daily/casual labourers for their farm works and employed wageworkers for the same work or herding and house services are less likely to engage in income diversification than others.

Livestock ownership is rewarding in the rural areas. Thus the greater the number of livestock a household possesses the less the interest in diversifying income. As the univariate result indicates, this hypothesis is supported by the regional sample data.

Fertiliser usage and food balance, indicators of a household's economic status, are found to be important determinants of the probability of involvement in different RNFAs. Farm households who are users of fertiliser are less likely interested in off-farm activities when compared with nonusers. The data also revealed that rural households who are net buyers of grain (for consumption) are more likely involved in income diversification than those better off in agricultural output.

As far as main employment status of heads is concerned, households who are headed by farmers working for other farm households are more likely to practice non-farm activities compared with employer households.

The univariate results described above show that non-farm activities are mostly practiced by economically weak households and are serving as a means of survival.

Table 1: PREDICTORS OF THE PROBABILITY OF INVOLVEMENT IN RNFAs

(Univariate Logistic Regression Results (at $\alpha=.15$))

Variable	COeff.(b)	S.E(b)	Wald	df	Sig	R	Exp(b)
SEX OF HEAD(MALE)	-.4644	.1465	10.0446	1	.0015	-.0698	.6285
MARITALS TATUS OF HEAD(MARRIED)	-.5302	.1382	14.7170	1	.0001	-.0878	.5885
RELIGION OF HEAD(ORTHODOX)	-.3293	.1442	5.2130	1	.0224	-.0441	.7194
HOUSEHOLD HEAD AGE			14.0451	3	.0028	.0698	
31-40	.5829	.1668	12.2166	1	.0005	.0787	1.7912
41-50	.5057	.1753	8.3242	1	.0039	.0619	1.6581
>50	.2903	.1772	2.6839	1	.1014	.0203	1.3368
MAI EMP.STATUS OF HEAD			5.6381	2	.0597	.0315	
OWN ACC. WORKER	.276	.1775	2.4196	1	.1198	.0159	1.3115
EMPLOYEE	.5389	.2269	5.6378	1	.0176	.0469	1.7141
RENT IN FARMLAND(YES)	-.3425	.1252	7.4766	1	.0063	-.0576	.7100
SHORTAGE OF LABOUR(YES)	-.4273	.1511	7.9967	1	.0047	-.0603	.6523
EMPLOY PERMA.LABOURER(YES)	-1.3383	.2110	40.2372	1	.0000	-.1522	.2623
INVOL.IN CROP PROD.(ENGAGED)	-.2991	.1637	3.3384	1	.0677	-.0285	.7415
FERTILISER USE(YES)	-.3183	.1949	2.6675	1	.1024	-.0201	.7274
AGRICULTURAL INCOME			23.1737	3	.0000	.1020	
3.01-6 QT	-.2339	.1623	2.0763	1	.1496	-.0068	.7914
6.01-11QT	-.4107	.1613	6.4853	1	.0109	-.0521	.6632
>=11QT	-.8085	.1720	22.0958	1	.0000	-.1103	.4455
NUMBER OF LIVESTOCK			7.8045	3	.0502	.0331	
3-6	.0602	.1574	.1460	1	.7024	.0000	1.0620
7-11	-.1039	.1642	.4003	1	.5269	.0000	.9013
>11	-.4035	.1739	5.3859	1	.0203	-.0453	.6680
FOOD BALANCE PER MEMBER			32.3203	3	.0000	.1262	
-2.2624-0.8QT	-.2007	.1592	1.5888	1	.2075	.0000	.8182
-0.79-0QT	-.7151	.1684	18.0365	1	.0000	-.0985	.4891
>0 QT	-.8240	.1734	22.5726	1	.0000	-.1116	.4387

4.1.2 Multivariate Results

Almost all of the predictors that are found to have marked association with participation in RNFAs, in the univariate case, have come out as important influencing factors at the multivariate level.

The multivariate procedure captures the effect of a predictor variable, that may or may not be weakly associated with the response variable at the univariate level. This enable to see the influence of each predictor on the probability of involvement in non-agricultural activities given other regressors are already in the model. This is, in fact, an advantage of the multivariate approach over its univariate counterpart.

There is no variable selection procedure that is superior than another, but either of them can be employed depending on the interest of the analyst. Based on parsimony and interpretation of results the **backward stepwise** procedure (with the **likelihood ratio test statistic** for removal) isemployed for this analysis. The output of this procedure is displayed in Table 2.

The results listed under the column labeled with **coeff (b)** are the maximum likelihood estimates, **b**, of the parameter vector β of the logistic regression model (i.e., they are the coefficients of the corresponding predictors, shown under the “**variable**” column, in the fitted multiple logistic regression model). A negative (positive) value down this column indicates that a particular regressor is inversely (directly) related to the probability of participation in income diversification. A value in this column shows the change in log-odds due to a change in levels in a regressor. **Exp(b)** shows the factor by which the odds of involvement in RNFAs change as a result of a change in level of a particular predictor. A negative (or positive) b_i , resulting

in $\exp(b_i) < 1$ (or $\exp(b_i) > 1$) implies a decrease (or increase) in the odds of involvement of rural households in non-farm activities.

The **S.E (b)** column, in Table 2, contains the standard errors of the estimators of the parameter β in the multiple logistic regression model.

The columns headed with “**Wald**” and “**Sig**” serve for one and the same purpose – to decide on the importance of a predictor in explaining the probability of involvement in RNFAs. The Wald statistic, which is the square of the ratio of **b** to **S.E (b)**, compared to its asymptotic chi-square value with the indicated degrees of freedom (labeled with **df** in Table 2) helps to test the null hypothesis that $\beta_i = 0$ against the alternative $\beta_i \neq 0$. If **sig** exceeds a specified level of significance, $\alpha = 0.1$, then we conclude that the corresponding predictor is unimportant in explaining the probability of involvement given that other predictors are in the model.

The values under the heading **R** show the partial correlations between the probability of involvement in RNFAs and each of the regressors in the model.

The goodness of the fitted model can be assessed via the classification table (Appendix I) or the goodness-of-fit statistics in appendix II. The model chi-square value, which is the difference between **-2LL** for the fitted model with the constant term only and **-2LL** for the fitted model including the constant with 21 degrees of freedom shows that all the variables included in the final model are significant. The goodness-of-fit value or the **-2LL** for the fitted model with the constant is less than $\chi^2_{(n-p)}(0.05) = 1388.88$ signifying that the multiple logistic regression model is fit. The classification table ascertains this result.

The different case-wise plots of residuals, the Half-Normal probability plot of standardised residuals and collinearity measures (condition indices and variance inflation factors (VIFs))

and leverage values were assessed to see the adequacy of the fitted model. The influence of each observation on the model parameter estimates was checked using DfBETAs values.

Extreme observations and high leverage values that influence the overall fit of the model and the individual estimates were observed both from the case-wise residual plots and Half-Normal probability plot of standardised residuals. Highly influential observations were also detected on the case-wise plots of leverage values and DfBETAs.

As regards near linear dependencies in the predictor variables, the computed condition indices (the maximum of them is less than 56) and VIFs (the largest of all smaller than 5) assures that the regressors are free from multicollinearity.

The observed inadequacy of the fitted model was corrected by removing cases with standardised residuals, leverage values, DfBETAs, and cook's distance exceeding $|\pm 3|$, 0.04, 0.05 and 0.1 respectively (see Appendix III).

Table 2. Multivariate Logistic Regression Results (at $\alpha=10\%$)
(IMPORTANT PREDICTORS OF THE PROBABILITY OF INVOLVEMENT IN RNFA's)

VARIABLE	Coeff(b)	S.E(b)	Wald	df	Sig	R	Exp(b)
SEX OF HEAD(MALE)	-.5500	.1684	10.6620	1	.0011	-.0743	.5770
MARITAL STATUS(MARRIED)	-.4439	.1645	7.2830	1	.0070	-.0580	.6415
AGE OF HEAD			17.5213	3	.0006	.0857	
31-40	.7522	.1919	15.3650	1	.0001	.0923	2.1216
41-50	.6920	.2038	11.5305	1	.0007	.0780	1.9977
ABOVE 50	.4523	.2069	4.7768	1	.0288	.0421	1.5719
RELIGION (ORTHODOX)	-.4347	.1660	6.8588	1	.0088	-.0557	.6475
HOUSEHOLD SIZE(≥ 6)	.4520	.2073	4.7539	1	.0292	.0419	1.5715
MAIN EMP.STATUS OF HEAD			8.9985	2	.0111	.0565	
OWN ACC. WORKER	-.3392	.2153	2.4811	1	.1152	-.0175	.7124
EMPLOYEE	.2093	.2711	.5959	1	.4401	.0000	1.2328
AGRICULTURAL YIELD			8.2969	3	.0403	.0383	
3.01-6 QT	-.3867	.2171	3.1732	1	.0749	-.0274	.6793
6.01-11 QT	-.3560	.2187	2.6494	1	.1036	-.0204	.7005
> 11 QT	-.6863	.2390	8.2461	1	.0041	-.0631	.5034
FOOD BALANCE PER MEMBER			28.3875	3	.0000	.1195	
-2.2624- -0.8 QT	-.0993	.1777	.3126	1	.5761	.0000	.9054
-0.79-0 QT	-.6521	.1908	11.6738	1	.0006	-.0785	.5210
> 0 QT	-.9121	.2017	20.4412	1	.0000	-.1084	.4017
EMPLOY PER.LABOURER(YES)	-2.6648	.4044	43.4111	1	.0000	-.1625	.0696
INVO.IN CROP PROD.(ENGAGED)	-.4436	.2404	4.2541	1	.0392	.0379	.6417
LABOUR SHORTAGE(YES)	-.4730	.2052	5.3135	1	.0212	-.0460	.6231
DAILY LAB. HIRED(2-81)	.7974	.3527	5.1120	1	.0238	.0445	2.2197
ADDIT.INCOME SOURCE(YES)	-.5639	.1459	14.9381	1	.0001	-.0908	.5690
RENT IN FARMLAND(YES)	-.3126	.1470	4.5213	1	.0335	-.0401	.7316
CONSTANT	-1.4511	.2773	27.3799	1	.0000		

4.2 DISCUSSION

We can hypothesize on how involvement in income diversification can be determined by different factors. However, the fact that these hypotheses are not supported by the empirical data may not lead to the conclusion that the data could be wrong, because facts about RNFAs could be obscured due to the influence of socio-economic and cultural situation of the area under study. That is to say, studies on the determinants of involvement in non-farm activities ought to consider the overall economic, social and cultural situations of areas covered by a study.

The results of this study are interpreted by paying attention to religious, cultural and economic situation of the rural population of the Amhara region. As data about the influence of culture on involvement in non-farm activities was not collected by the data source and due to absence of published documents regarding the issue, the opinion of individuals familiar with the region is used while interpreting some of the model results.

The univariate and multivariate logistic regression results indicate that a number of factors determine involvement in RNFAs in the rural Amhara.

The demographic variables age, sex, religion, marital status and household size significantly determine the probability of involvement in non-agricultural activities. It appears that age of household head has a positive influence on the probability of involvement in RNFAs. However, a declining trend in participation in non-farm activities is observed the older the head. The odds ratio shows that the odds of involvement in rural non-farm activities increases by twofold and by 1.6 in households whose heads are in the age groups 31 to 40, 41 to 50 and above 50 respectively as compared with households whose heads' age lie between 18 and 30. This result supplements the finding by Mulat and Teferi in North Shoa. Tegegne (1996) also

revealed that farmers participating in income source diversification were younger than those involved in agriculture only in two cash crop potential/growing *weredas* of southern Ethiopia. The primary reason could be due to the fact that some activities, such as pottery, blacksmithery, and tanning, have been stigmatized in some parts of the region. On the other hand, given the cultural situation of the Amhara region, farmers are less likely to be engaged in non-farm activities in localities where their ancestors have lived unless they have completely lost hope on their mainstay, agriculture. The other reason could be that newly established households might strive for building sustainable economic reserve through diversifying their income as compared to long-established and experienced households. Moreover, experienced households may not venture into activities that are different from the old established traditional farming as their experience, under normal circumstances, is traditional agriculture.

Sex of household heads is also another differential factor that influences involvement in non-agricultural activities. The result of the sample data shows that the odds of being involved in RNFA for female-headed households increases by more than 57 percent as opposed to male-headed ones. The univariate logistic regression finding has shown a similar result. This reveals that female-headed households are more likely involved in different income generating activities. The result is actually a typical one confirming the anticipation that most non-agricultural activities are women's business. Similar results were found by MOLSA, at the national level, and by Mulat and Teferi, in North Shoa.

A lot of reasons could be cited why this is so. In the first place, some activities, such as brewing and sale of local drinks (*araki* and *tella*) and food preparation, are generally considered to be the sole tasks of women (Mulat and Teferi, 1996). Pottery and spinning, accounting for 13 and 9 percent, respectively of engagement in the region (MOLSA, 1997), are predominantly women's activities and their participation in petty trades is also considerable. Usually, women

involved in RNFAs are single, divorced or widowed. This allows them to participate in income diversification freely. Though it is not documented there is a general consensus that most rural female-headed households are without sufficient plot of farmland.

The probability of involvement in RNFAs is also determined, both individually and jointly, by marital status of household heads. The odds of diversifying income sources in single, divorced, or widowed households is about one-half times higher than the value for their married counterparts. The reason for this is probably related to being disadvantaged with respect to farmland allotment and the household's overall economic status. Usually, in rural areas newcomers, like the households mentioned above, are less likely to get adequate farmland. Cultural influences are not also uncommon. For instance, the husband is hardly willing to let his wife undertake non-agricultural activities, especially the stigmatized ones.

As regards religion, followers of the Orthodox Church are less likely to get involved in RNFAs. The odds of involvement in income diversification for followers of Orthodox Church decreases by more than 64 percent as opposed to Muslims, Protestants or Catholics.

As studies on non-farm activities in Africa and in particular in Ethiopia are scanty and the available ones are not specific enough and, concerned with some characteristics of the sector only, it is often difficult to compare results regarding some influencing factors of RNFAs. However, MOLSA's national result is acting as a model for comparison. According to MOLSA (1997), of the 24 percent households engaged in weaving in the country the majority (41 percent) were in rural Amhara. RRC (1994) reported that this activity is dominated by Muslims. On the other hand, trade in consumer goods, the largest share of trade activities in the rural areas (MOLSA, 1997), is usually undertaken by Muslim petty traders. This in part ascertains that the result of the empirical data is not without truth.

Population growth, which is becoming a pressing issue, is felt in the Amhara region where population is growing at about 2.86 percent annually as of Oct.1994 (CSA, 1998). Obviously, this does have a negative effect on development especially in areas where income sources are limited. It is necessary to assess whether increasing household size is a cause for income source diversification.

The multivariate logistic regression result shows that family size determines households' involvement in non-farm activities. It is found out that the odds of involvement in diversifying income sources increases by about 57 percent for households with six or more members when compared with those having five or less (**Odds Ratio** =1.5715). An independent t-test result revealed that the two groups (involved in non-farm and agriculture only) are, on the average, endowed with the same family member ($p = 0.839$). The same test revealed that households involved in RNFAs have lower yield than those involved in agriculture only ($p=0.002$). This implies that agricultural yield is a decisive factor for a household's decision on income diversification.

Different studies have come up with different findings on this issue. Mulat and Teferi (1996) and Tegegne (1996) found out that family size is not a differential factor for farm income. On the other hand, in his comparison of the two groups Tegegne (1996) indicated that households who participated in non-farm activities were found to be less endowed in terms of family labour.

Some endowment and related factors are also found to be significant determinants of participation in RNFAs. These are: agricultural yield, food balance per family member, the number of labourers hired for farm work, renting farmland, involvement in crop production, presence

of additional income source, shortage of labour and whether a household employed wage worker for farm work, herding or household service.

As has been indicated by Mulat and Teferi (1996), households are less likely interested in non-farm employment as long as they are satisfied with their agricultural yield. In fact, the regression results suggest that participation in RNFAs generally decreases as households' yield increases. The odds ratio (Table 2) indicates that the odds of income diversification for households who managed to produce 3.01 to 6, and 11 or more quintals of agricultural output decreases by about 68 and 50 percent, respectively relative to those who could harvest up to three quintals of agricultural yield. Mulat and Teferi (1996) found a similar result in North Shoa. That is, farm households with less dependable agricultural base are highly likely to diversify their income. The finding is what can be anticipated given the agrarian nature, the cultural barrier on some non-farm activities and the level of skill/knowledge of the rural population.

The amount of grain a household sells or buys for consumption tells something about engagement in RNFAs. It can be inferred from the result (Table 2) that the probability of involvement in RNFAs, in general, increases as households become food insufficient. Actually, the levels of the variable " food balance per household member" indicate that a large number of rural households were net buyers of grain in the reference period of the survey. The result is in line with our expectation, under the existing situation in the region, and this could be due to the fact that once households are satisfied with the food requirement of the family, they will be less interested in income diversification.

Hiring daily labourers or wage workers for farm work, or herding and household services, which is mostly a characteristic of wealthy households, has something to do with income

source diversification. The odds of participation in RNFAs for households who hired 2 to 81 labourers is about 2.2 times higher than those who hired one or none. This shows that rural households are tempted to take up RNFAs for wealth accumulation, or households who performed their farm activities earlier spare their time on non-agricultural activities.

Shortage of labour for farm work is negatively related with involvement in RNFAs. The result suggests that participation is lower in those households who fall short of labour. In fact, their odds of participation decreases by more than 60 percent. This may be due to the fact that non-employer households might have no time to engage in non-farm activities.

Some households rent-in farmland usually from female-headed, inactive and economically weak households. These households are mostly newly established ones, immigrants, or those who might have not enough farmland. However, we do not have enough evidence to categorise these households as poor or rich echelons of the population. But from experience the majority of renting households has been observed to have at least their own oxen, cows, or horses and plough/cart that indicate wealth of a household.

Whether a household rent-in farmland influences involvement in non-farm activities. The corresponding odds ratio in Table 2 suggest that households who did not rent-in plots of land have a better chance (73%) of diversifying their income sources. This is partly due to the reason that households who rent-in farmland have more farm resources and hence are less attracted to RNFAs.

Very few rural households derive additional income from remittances, animal rental, sale of different agricultural products (such as milk and milk products), and others. Besides the small amount, income from these sources may not motivate households to involve in off-farm activities. As a matter of fact, the contribution of the variable "additional income" in explaining

involvement in RNFAs is found insignificant. Nonetheless, whether a household has additional income source(s) or not influences involvement in RNFAs. Rural households who have additional income source are less likely to participate in different non-farm activities. Because, as we often see, whatever income a household derives it will be invested on traditional agriculture.

4.3 Identification of Wealth Groups in RNFAs

Identification of household groups involved in RNFAs is informative to distinguish differences in the use of income from these activities for different livelihood strategies. This type of grouping is helpful to identify the distribution and choice of RNFAs that a farm household occupies.

Wealth indicators such as farm size, grain yield, livestock number, plough/cart and number of daily labourers a household hired for farm/non-farm activities can be used as central parameters for grouping peasants into different wealth groups based on wealth index created (Møller, 1998). Some of these indicators vary depending on the type of farming practiced and crop produced in a region. Moreover, the number of indicators used for classification is not fixed.

Total agricultural yield (in quintal), farm size, livestock number, annual income (from the sale of farm/non-farm products) and the number of daily labourers a household hired could be used in classifying RNFAs actors in the Amhara region.

The wealth indicators are most often correlated, but in many cases farm households differ from average at one or more of the indicator variables.

The formation of the three wealth groups and hence identification of the types of activities by each group, which has some subjectivity, is described in the following way.

The values of each indicator variable will be assigned a value of 0, 2, or 4 (the numbers are chosen for simplicity) depending on the size of that value relative to the first and third quartiles of the indicators (Table 3). The sum of points assigned for a household for all the five variables will not exceed 20.

Table 3 Indicators for Wealth Grouping

Indicator Variable	Points Assigned		
	0	2	4
Total Agricultural yield (in qt)	0- 2.8750	2.8751-11	> 11
Farm size (in ha)	0- 0.3025	0.3026 - 1.5	> 1.5
Livestock	0	1-4	> 4
Annual income (in Birr)	0	0.001-184.3750	> 184.3750
Hire of day labourer	0	1-2	≥ 3

Source: Adapted from Møller (1998)

The numbers in the table (except the last row) are the first and third quartiles of the corresponding indicator variables.

From the distribution of observations (Table 3) three wealth groups (“poor”, “medium”, “rich”) can be derived .We categorise a household as “poor” if the resulting sum of points for a given household lies in the range 0 to 6, or as “medium” if the resulting sum fall in 8 to 12 or as “rich” otherwise (Table 4).

Table 4. Distribution of households in RNFAs based on Wealth Indicators

	Wealth Groups		
	"Poor"	"Medium"	"Rich"
Indexpoint	0- 6	8- 12	14-20
Distribution(%)	30.4	54.6	15.0

Based on the above index points activities performed by households involved in RNFAs are identified.

Tables 5, 6, and 7 below show the different non-farm activities that were performed as main activities by the different wealth groups in the rural parts of the Amhara region.

Accordingly, weaving was a dominant RNFA in the region that was undertaken by relatively better off households. The fact that all rural households are users of the traditional Ethiopian clothes, *gabi* and others, easiness to be done even at night, and when raining and its low initial capital make the activity accessible. The payment made to the service mostly in cash and its increase from time to time enables weavers to supplement their agriculture.

Construction (mainly masonry) was the main activity of households earning better non-farm income and/or agricultural return. Masons generate a substantial amount of money especially during the construction of schools, clinics, churches and other infrastructures nearby.

Of the handicraft and related main activities blacksmithing was practiced by medium wealth groups of households. As they are the ones who repair farm instruments, the income derived by blacksmiths is not negligible. Pottery was the third main activity undertaken by poor and medium groups of households. According to the grouping, potters are not disadvantaged compared with the other households engaged in income diversification.

In studying the distribution of handicraft activities (Table 5), it appears that poor RNFA actors were engaged in low paying and less technical activities whereas medium and relatively rich households were involved in pottery, carpentry, weaving, blacksmithing and masonry which require skill, and large initial capital.

The wealth grouping of households involved in RNFAs revealed that households who were economically weak were engaged, as anticipated, in the sale of food and drink and petty trades (Tables 6 and 7).

Crops, livestock, butter/honey and most commodity trades were the activities of resourceful households. These activities are relatively highly profitable and their potential in supplementing meager agricultural yield is considerable.

Table 5. ENGAGEMENT ON HANDICRAFT AND RELATED ACTIVITIES

HANDICRAFT and RELATED ACTIVITIES	WEALTH GROUP			TOTAL
	"POOR" 0 – 6	"MEDIUM" 8 – 12	"RICH" 14 – 20	
WEAVING	7(6.4)	80(72.7)	23(20.90)	110
SPINNING	10(58.8)	6(35.3)	1(5.9)	17
POTTERY	9(29)	16(51.6)	6(19.4)	31
BLACKSMITHING	4(11.8)	24(70.5)	6(17.6)	34
TANNERY	1(25)	3(750)	-	4
CARPET MAKING	-	-	1(100.0)	1
BAMBOO WORKS	4(26.7)	11(73.3)	-	15
'ZENBIL' AND THE LIKE	-	1(100.0)	-	1
'SIFET' WORKS	4(36.4)	5(45.5)	2(18.2)	11
TAILORING	3(30)	2(20.0)	5(50.0)	10
WOODEN H.H MATERIAL	-	4(66.7)	2(33.3)	6
CONSTRUCTION	1(4.8)	12(57.10)	8(38.1)	21
MINING	1(100.0)	-	-	1
SWEATER	5(35.6)	4(28.6)	5(35.7)	14
OTHERS	5(23.8)	11(52.4)	5(23.80)	21
TOTAL	54 (18.2)	179 (60.3)	64 (21.5)	297

Table 6. ENGAGEMENT ON TRADE ACTIVITIES

TRADE ACTIVITY	WEALTH GROUP			TOTAL
	"POOR" 0-6	"MEDIUM" 8-12	"RICH" 14-20	
FOOD	2 (22.2)	6 (66.7)	1 (11.1)	9
DRINK	22 (36.7)	29 (48.3)	9 (15.0)	60
TRADE IN CROPS	1 (6.3)	12 (75.0)	3 (18.8)	16
TRADE IN LIVE STOCK	-	1 (50.0)	1 (50.0)	2
BUTTER/HONEY TRADE	-	3 (60.0)	2 (40.0)	5
COMMODITY TRADE	5 (21.7)	12 (52.2)	6 (26.1)	23
CHICKEN/EGG TRADE	-	1 (100.0)	-	1
OTHER TRADE	10 (66.7)	3 (20.0)	2 (13.3)	15
GRAIN MILL	1 (33.3)	2 (66.7)	-	3
OIL MILL	-	1 (100.0)	-	1
TOTAL	41 (30.4)	70 (51.9)	24 (17.8)	135

Table 7. ENGAGEMENT ON FOOD AND DRINK ACTIVITIES

FOOD and DRINK TRADE ACTIVITY	WEALTH GROUP			TOTAL
	"POOR" 0-6	"MEDIUM" 8-12	"RICH" 14-20	
FOOD	1 (20.0)	3 (60.0)	1 (20.0)	5
DRINK	23 (34.8)	34 (51.5)	9 (13.6)	66
TOTAL	24 (33.8)	37 (52.1)	10 (14.1)	71 (100.0)

.The numbers in parentheses in Tables 5, 6, and 7 are row percentages.

Chapter V

CONCLUSION AND RECOMMENDATION

The reviewed studies in Ethiopia, Kenya, Sudan and other few African countries indicate that rainfed agricultural return is declining from time to time. Rural households overcome this crisis through combining agriculture with one or more non-farm activities. African households need to diversify their income as the best livelihood strategy.

The small average landholding size (about 1 hectare) in the Amhara region is an indication of considerable pressure on farmland. This together with low productivity of the land, due to man-made and natural calamities, is expected to force rural households to move into activities beyond traditional agriculture.

However, the existing level of involvement in RNFAs in the rural areas of the region leaves much to be desired when viewed from the point of view of food security and overall environmental degradation. Though some activities, like pottery, blacksmithing and tanning, have been performed since many years ago, the overall involvement, unlike some areas in Africa, is very low.

Income diversification is mostly practiced by a small segment of the rural population. As the results of this study show, some rural households especially those having less dependable resource in the agricultural sector are taking steps to supplement their family requirements and subsidise the declining agricultural return through different non-agricultural activities.

The logistic regression results show that the odds of participation is higher in female headed, relatively newly established and Muslim households. These, in one way or the other, indicate the presence of cultural barrier on some non-farm activities. In fact, blacksmithing, tanning

and pottery, despite their contribution in the day-to-day activities of rural Amhara, are not expanded due to the low status accorded to handicraft activities.

The finding of this study shows that rural households with larger family size tend to participate in RNFAs than those having less household members. This is actually an indication that largeness of a household is pushing households to look for alternative means of livelihood. The finding does not, however, lead to the conclusion that households involved in RNFAs are those endowed with more usual household members. In fact, an independent t-test revealed that farm households involved in RNFAs and agriculture only are endowed with nearly equal average family size.

The corresponding odds ratios for the endowment variables such as agricultural yield and food balance reflect that the largest proportion of households involved in income diversification has small or no farm land. The demographic variables sex and marital status of head and main employment status of head partly reflect the situation. The findings on agricultural yield, food balance and main employment status of heads further indicate that participation in non-farm employment is induced by low agricultural income or assets. Most better-off households in the region may not engage in non-farm activities presumably due to the presupposition that income from these activities is lower than income from agriculture.

Hiring more labour enables households to complete their farm activities earlier. This may allow in involving in income diversification if they wish to do so.

Households having employed wage worker for farm work, herding and household service, and those who faced labour shortage for their farm activities are less likely to engage in activities other than agriculture. However, the results also show that these households will have

a greater tendency to participate in these activities if they managed to hire a larger number of daily labourers for their farm works.

Lack of credit facilities was reported as a major problem by 39.6, 42.9 and 19.2 percent of households involved in food and drink, trade and handicraft and related activities, respectively. Furthermore, about 12.9 percent of households undertaking handicraft and related activities faced high purchase price of raw materials; low sales price was a problem of 23.8, 15.1 and 10.2 percent of households engaged in handicraft, trade and food and drink activities. It is therefore essential to expand the scope of the available credit schemes so as to make RNFA's sustainable in fulfilling household requirements. But the schemes should be implemented in such a way that some segment of the population will be willingly shifting to non-agricultural sectors. It is also advisable to encourage better-off households to engage in different non-farm activities to accumulate good reserve to combat any adverse effect on rain-fed agriculture. It is known that some traditional handicraft activities have been stigmatized though the outputs of these activities are the dominant tools in the day-to-day activities of the rural households. This outlook has to be reversed, by launching different awareness creation programmes, so as to make the sector attractive.

The data show that almost all of (99.9 percent) the households engaged in handicraft and related activities did not obtain training and other assistance from governmental and/or non-governmental institutions. The role of non-farm activities in reducing land fragmentation and supplementing meager agricultural revenue will be high provided that concerned governmental organisations, the Agency for Handicraft and Small scale Industries (DAHASII), Rural Technology Promotion Division of the Bureau of Agriculture (RTPD) and the Community Skill Development Centers of the Bureau of Education (CSDCs) are strengthened, reoriented and coordinated to address the felt-needs of rural population of the region. The contribution of

NGOs should not be overlooked in this regard. It should be emphasised that any intervention programme in the sector should consider the cultural, religious and socio- economic situation besides the full participation of the farmers themselves.

Most non-farm activities are so traditional that they may not enable actors to derive sufficient income out of it. Moreover, the activities are being undertaken by farmers with no or inadequate skill or knowledge. And therefore, concerned institutions need to emphasise on practical and market oriented trainings and modification of the mode of productions of handicraft activities.

The data revealed that households involved in RNFAs do have small number of farmlands and derived small amount of agricultural yield as compared with those engaged in agriculture only. The available evidences show that shortage of farmland forced most RNFAs to take up different alternative income sources. This suggests that households involved in non-agricultural activities have to get support to enhance their agricultural productivity as the non-farm sector, at its present stage, alone is not sufficient to lead the life of a household.

Some non-agricultural activities such as food and drink and trade are becoming increasingly important in peasant economy, and hence national and regional researches should give due emphasis to the sector. The complementary nature of agriculture and the sector has to be assessed in future studies. Moreover, in-depth investigations on the influence of culture on each of the sub-sectors of RNFAs need to be investigated.

REFERENCES

1. Agresti, A. (1996). *An introduction to Categorical Data Analysis*. John Wiley & Sons Inc., New York.
2. Bagachawa, M.S.D. (1997). The Rural Informal Sector in Tanzania. **In:** *Farewell to Farms: Deagrarianisation and Employment in Africa* Bryceson, D.F. and Jamal, V. (eds.) pp.137-154. African Studies Centre, Leiden.
3. Berkvens, R.J.A (1997). *Backing two Horses: Interaction of Agriculture and Non-Agricultural Household Activities in Zimbabwean Communal Area*. Working Paper Vol. 24. African Studies Center, Leiden.
4. BOE (1997). *Annual Education Statistics of the Amhara Region*. Bahir Dar (unpublished).
5. BOPED (1997). *Socio-Economic Situation of the Amhara National Regional State- 1993-97* (Amharic Version), **Vol. I**. Bahir Dar (unpublished).
6. Bryceson, D.F. (1997). Deagrarianisation in Sub-Saharan Africa: Acknowledging the Inevitable. **In:** *Farewell to Farms: Deagrarianisation and Employment in Africa*. Bryceson, D.F. and Jamal, V. (eds.) pp.3-20. African Studies Centre, Leiden.
7. Bull, S.B., and Donner, A. (1987). The efficiency of Multinomial Logistic Regression Compared with Multiple Group Discriminant Analysis. *Journal of the American Statistical Association*, **82**(400): 1118-1122.
8. Central Statistical Authority (1992). *Rural Labour Force Survey (1987/88): Statistical Bulletin NO. 108*. CSA, Addis Ababa.

9. _____ (1998). *The 1994 Population and Housing Census of Ethiopia: Results of the Amhara Region*, Vol.V. CSA, Addis Ababa.
10. _____ (1998). *Agricultural Sample Survey (1997/98): Report on Land Utilization*, Vol. IV. CSA, Addis Ababa.
11. _____ (1998). *Agricultural Sample Survey 1997/98: Report on Area and Production for Major Crops (Private Peasant Holdings, Meher Season)*, Vol.I. CSA, Addis Ababa.
12. Collet, D. (1991). *Modelling Binary Data*. Chapman & Hall, London.
13. Copas, J.B. (1988). Binary Regression Models for Contaminated Data. *Journal of Royal Statistical Association*, B50 (2): 225-265.
14. Desta Weldemariam (1990). *Determinants of Fertilizer Use in the North Western Ethiopia*. (Unpublished M.Sc. Thesis), Addis Ababa University.
15. Foeken, D.(1997). Urban Trajectories: Household Employment Patterns in Kenya's Coast Province. **In:** *Farewell to Farms: Deagrarianisation and Employment in Africa*. Bryceson, D.F.and Jamal, V. (eds.) pp.119-136. African Studies Centre, Leiden.
16. Gaidzanwa, R.B. (1997). Non-Farm Activities and Gender in Zimbabwe. **In:** *Farewell to Farms: Deagrarianisation and Employment in Africa*. Bryceson, D.F.and Jamal, V. (eds.) pp.157-166. African Studies Centre, Leiden.
17. Hauck, W.W and Donner, A. (1977). Wald's Test as Applied to Hypotheses in Logit Analysis. *Journal of American Statistical Association*, 72: 851-853.
18. Hosmer, D.W. and Lemeshow, S. (1982). A review of Goodness-of-fit Statistics for use in the Development of Logistic Regression Models. *American Journal of Epidemiology*, 115(1): 92-106.

19. Hosmer, D.W. and Lemeshow, S. (1989). *Applied Logistic Regression*. John Wiley & Sons, Inc. New York.
20. Ibrahim, H.B. (1997). Coping with Famine and Poverty: The Dynamics of Non-Agricultural Rural Employment in Darfur, Sudan. **In:** *Farewell to Farms: Deagrarianisation and Employment in Africa*. Bryceson, D.F. and Jamal, V. (eds.) pp.23-40. African Studies Centre, Leiden.
21. Iliya, M.A and Swindel, K.(1997). Winners and Losers: Household Fortunes in the Urban Peripheries of Northern Nigeria. **In:** *Farewell to Farms: Deagrarianisation and Employment in Africa*. Bryceson, D.F. and Jamal, V. (eds.) pp.85-100. African Studies Centre, Leiden.
22. ILO, UNDP, UNIDO and The Government of Netherlands (1988). *Development of Rural Small Industrial Enterprise: Lesson from experience*. Vienna.
23. Jambiya, G. (1998). *The Dynamics of Population, Land Scarcity, Agriculture and Non-Agricultural Activities: West Usambara Mountains, Lushoto District, Tanzania*. Working Paper, Vol.28. African Studies Center, Leiden.
24. Jennings, D.E. (1986). Outliers and Residual Distributions in Logistic Regression. *Journal of the American Statistical Association*, **81**(396): 987-990.
25. Landwehr, J.M., Pregibon, D., and Shoemaker, A.C.(1984). Graphical Methods for Assessing Logistic Regression Models. *Journal of the American Statistical Association*, **79**(385): 61-71.
26. McCullah, P. and Nelder, J.A. (1989). *Generalised Linear Models*. Chapman & Hall, London.

27. Meagher, K. and Mustapha, A.R. (1997). Not by Farming Alone: The role of Non-farm Incomes in Rural Hausaland. **In:** *Farewell to Farms: Deagrarianisation and Employment in Africa*. Bryceson, D.F. and Jamal, V. (eds.) pp.63-84. African Studies Centre, Leiden.
28. Møller, M.R. (1998). *The Changing Roles of Rural Non-Agricultural Activities in the Livelihoods of Nigerien Peasants*. Working Paper 98.9. Center for Development Research, Denmark.
29. MOLSA (1997). *Agricultural Wage Employment and Rural Non-Farm Employment in Ethiopia: Survey Results*. Addis Ababa, Ethiopia.
30. Montgomery, D.C. and Peck, E.A (1992). *Introduction to Linear Regression Analysis*. John Wiley & Sons Inc., New York.
31. Mulat Demeke & Teferi Regassa (1996). Non-Farm Activities in Ethiopia: The case of North Shoa. **In:** *The Ethiopian Economy, Poverty & Poverty Alleviation. Proceedings of the 5th Annual Conference on the Ethiopian Economy*. Bereket Kebede & Mekonnen Tadesse (eds.). pp.125 -162, Addis Ababa.
32. Mung'ong'o, C.G. (1998). *Coming Full Circle: Agriculture, Non-farm Activities and the Resurgence of out-migration in Njombe District, Tanzania*. Working Paper Vol.26. African Studies Center, Leiden
33. Rao, C.R. (1973). *Linear Statistical Inference and its Applications*. 2nd Ed. John Wiley & Sons Inc., India.
34. Reardon, T., Crawford, E. and Kelly, V. (1994). Links between Non-Farm Income and Farm Investment in African Households: Adding the Capital Market Perspective. *American Journal of Agricultural Economics*, **76**: 1172-1176.

35. Reardon, T., Delgado, C. and Malton, P. (1992). Determinants and Effects of Income Diversification amongst Farm Households in Burkinafaso. *The Journal of Development Studies*, **28**(2): 264-296.
36. RRC (1994). *Secondary Economic Activities* (unpublished). Addis Ababa, Ethiopia.
37. Rubin, D.B. (1984). Assessing the Fit of Logistic Regressions using the Implied Discriminant Analysis. *Journal of American Statistical Association*, **79**(385): 79-80.
38. Sincich, T. (1993). *Statistics by Example*. Macmillan Publishing Co., USA.
39. SPSS Institute Inc. (1994). *SPSS Advanced Statistics*, Version 6.1. SPSS Institute Inc., Chicago.
40. Tegegne Gebre Egziabher (1995). An assessment of Ethiopian's Agricultural Land Resources. **In:** *Ethiopian Agriculture: Problems of Transition. Proceedings of the Fourth Annual Conference on the Ethiopian Economy*. Dejene Aredo and Mulat Demeke (eds.), pp.22-39.
41. Tegegne Gebre Egziabher (1996). The Impact and Determinants of Rural Non-Farm Activities: The case of Damotegale and Kachabira *Weredas* in Ethiopia. *Ethiopian Journal of Development Research*, **18**(1): 49-84.
42. UN (1997). *The Least Developed Countries: 1997 Report*. New York.
43. Van Dijk, M.P. (1997). Economic Activities of the Poor in Accra. **In:** *Farewell to Farms: Deagrarianisation and Employment in Africa*. Bryceson, D.F. and Jamal, V. (eds.) pp.101-115. African Studies Centre, Leiden.

APPENDICES

APPENDIX I. CLASSIFICATION TABLE

OBSERVED	CORRECT PREDICTION (%)	
	NOT INVOLVED IN RNFAs	73.02
	INVOLVED IN RNFAS	58.13
	OVERALL	68.75

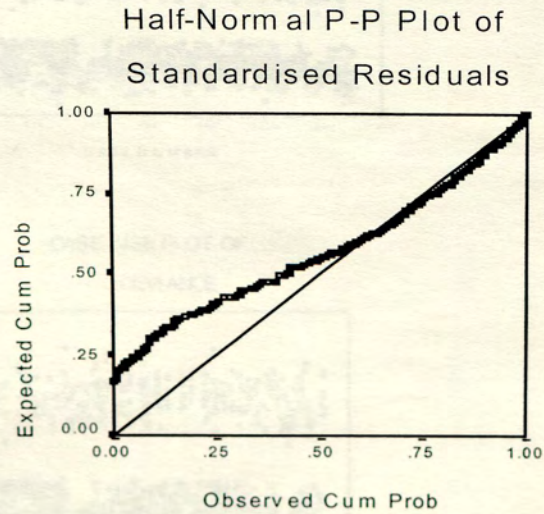
.The cut off value used is .35

APPENDIX II. GOODNESS OF FIT SUMMARY STATISTICS

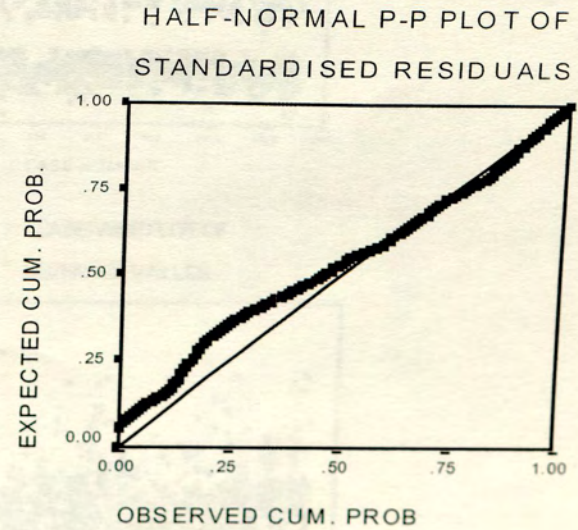
Statistics	Value
-2 Log Likelihood(Model with constant only)	1568.0986
-2 Log Likelihood(Fitted model with constant)	1339.710
Goodness of Fit	1171.862
Model Chi-square	228.388
Degrees of Freedom	21
Significance	.0000

APPENDIX III-1. HALF-NORMAL PLOTS OF STANDARDISED RESIDUALS
FOR DETECTION OF OUTLIERS

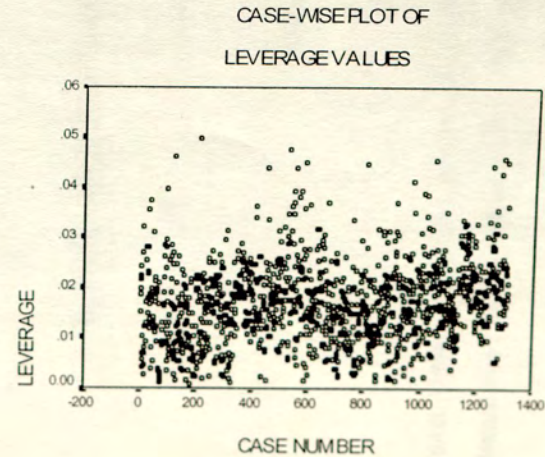
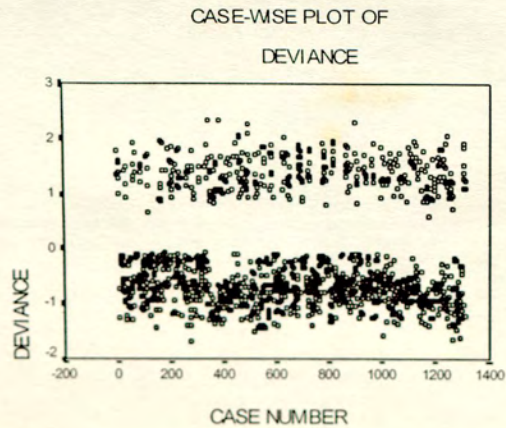
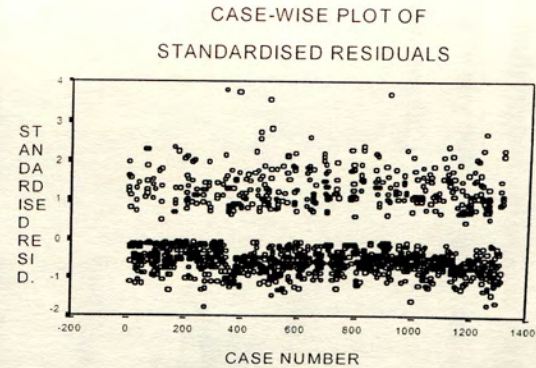
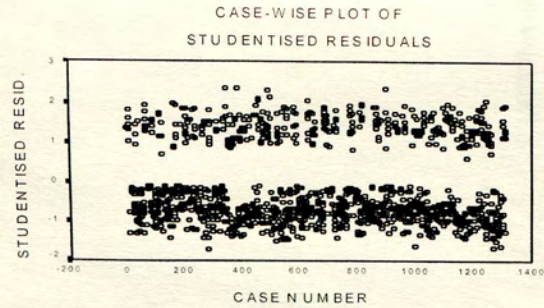
BEFORE DELETION OF OUTLIERS



AFTER DELETION OF OUTLIERS



APPENDIX III-2 CASE-WISE PLOTS OF RESIDUALS (AFTER OUTLIERS ARE REMOVED)



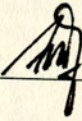
DECLARATION

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

Name

Abebaw Getachew

Signature



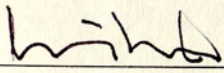
Place

Faculty of Science, Addis Ababa University

Date of Submission

June 2000

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



Dr. Eshetu Wencheke