Statistical Analysis of Determinants of Credit Loan Payment in Arba Minch Town, Ethiopia Using Logistic Regression Model

Belayneh Debasu Kelkay[*]**, Zelalem Esubalew**

Department of Statistics, College of Natural Sciences, Arba Minch University, Arba Minch, Ethiopia

Email address:
belaynehd@gmail.com (B. D. Kelkay)

To cite this article:
Belayneh Debasu Kelkay, Zelalem Esubalew. Statistical Analysis of Determinants of Credit Loan Payment in Arba Minch Town, Ethiopia Using Logistic Regression Model. *International Journal of Statistical Distributions and Applications*. Vol. 3, No. 1, 2017, pp. 7-12. doi: 10.11648/j.ijsd.20170301.12

Abstract: Ethiopia has one of the fastest-growing economies in the world and is Africa's second most populous country (IFC, 2012). In developing countries, like Ethiopia, financial resource is important input for continuous development. Most of the peoples living in under poverty line need wide range financial services for consumption, running their business and building assets. Due to lack of collateral, poor people in most cases have no credit access from Banks. Microfinance offers financial service such as loan, savings and micro insurance to the poor people either in individual or in a group basis to those people. The objective of this study isto investigate the factor that affects credit loan payment performance of urban people in ArbaMinchSeachasubcity. For this study the researcher would be used stratified sampling for proportional allocation because this study concerned on strata level.From 15336 total people live in this sub-town a sample of 95 person were randomly selected. The estimation results of the descriptive statistics and the logistic regression model show that age,sex, level of education, family size,attitude,awareness and number of times borrowed are important and significant factors that enhance using of credit loan activity. And the other variable is not significant factors that improve using of credit loan action like, marital status and religious.

Keywords: Credit Loan Payment, Logistic Regression, Secha Sub-city

1. Introduction

Ethiopia has one of the fastest-growing economies in the world and is Africa's second most populous country (IFC, 2012). Many properties owned by the government during the previous regime have now been privatized and are in the process of privatization. Almost 50% of Ethiopia's population is under the age of 18, and even though education enrollment at primary and tertiary level has increased significantly, job creation has not caught up with the increased output from educational institutes. The country must create hundreds of thousands of jobs every year just to keep up with population growth.The economy of Ethiopia is based on agriculture, which accounts for 46.3% of gross domestic product (GDP), 60% of exports, and 80% of total employment.

In Ethiopia, the formal and informal sectors co-exist, with differences in accessibility. Many financial institutions in developing countries provide financial services such as saving and credit to aid several households' enterprises. This is an effort in line with the millennium development goals which seeks to reduce poverty by 50 % by the year 2013. However, the sustainability and continuity of the financial institution to increase the volume of credit to stimulate the poverty reduction goal depend on the payment rate.

Several microfinance institutions (MFIS) have been established in Ethiopia and have been operating towards resolving the credit access problem of the poor particularly those who engage in little business (Befkadua, 2007). Generally the accessibility of a good financial service is considered as one of the engines of economic development. The establishment and expansion of financial service is also one of the instruments to break the vicious circle of poverty. Government of less developed countries has frequently practiced the policy of providing cheap credit to the agricultural sector through financial intermediaries. This cheap credit, it is hoped, will lower the dependence on the urban money lenders (Pinkie, 1998).

In developing countries, like Ethiopia, financial resource is important input for continuous development. Most of the peoples living in under poverty line need wide range financial services for consumption, running their business and building assets. Due to lack of collateral, poor people in most cases have no credit access from Banks. Microfinance offers financial service such as loan, savings and micro insurance to the poor people either in individual or in a group basis to those people. The achievement of microfinance activities in Ethiopia is mainly affected by the income of clients, which directly depends on the effectiveness of the small business of borrowers who live in urban areas.

Credit prevention is one of the principal components of urban development, which helps to attain rapid and sustainable growth of urban activities.. This study mainly involved on what are the source and the forms of credit practiced in the area. How the contribution of credit loan payment for the small households. This study is intended to deal with the following research questions:

- What are the factors that determine households likelihood of being credit constrained in Arba Minch town Secha sub city?
- What is the observation of customers regarding service freedom of formal financial institutions in the study area?
- What is the position of different wealth and sex groups' formal and informal credit?

Therefore, the main objective of this study is to investigate the factor that affects credit loan payment performance of peoples in Arba Minch town Sechasub-city by using binary logistic regression.

2. Methodology

2.1. Study Area

Arba Minch town is located about 454 km south of Addis Ababa and 275 km south of Hawassa. Its elevation range from 1200 meter above sea level at the Northern and 1320 above sea level at the southern end. The town has an average temperature 30°C and rain fall of 575mm, and the town has two sub – cities namely Sikela and Secha.

2.2. Target Population

The target population for this study would be all people from age 18 and above who live in secha sub-City.

2.3. Method of Data Collection

This study uses primary source of data using stratified sampling design method and obtained the data using a well-designed questionnaire.

2.4. Sample Size Determination

The sample sizes depend on level of significance (α), margin of error (d) and estimated of proportion (P&Q).Therefore, the formula to determine the sample size is:

$$n_o = \frac{z_{\alpha/2}^2 Pq}{d^2} = \frac{(1.96)^2(0.2)(0.8)}{(0.08)^2} = 96.04 = 96,$$ Where p = 0.2 and q = 0.8 are obtained by conducting pilot survey and α = 0.05 and d = 0.08

Since the value of $\frac{z_{\alpha/2}^2 Pq}{d^2}$ is greater than 5%, this study uses the value of n=$\frac{n_o}{1+\frac{n_o}{N}}$(Cochran, third edition, other wise use n=n_o).

Thus, n = $\frac{96}{1+\frac{96}{15336}}$= 95

$n_h = \frac{n}{N}N_h$Where n_h be sample drawn fromh^{th}stratum N_h be the population size in the h^{th}stratum.

N=N₁+N₂+N₃for three kebele's

Table 1. Stratum of Population Size.

Stratum	Population	Sample size to be drawn
Stratum-1	N1=total population of berekebele	$n_1 = \frac{N1*n}{N} = \frac{3872*95}{15336} = 24$
Stratum-2	N2=total population of doysakebele	$n_2 = \frac{N2*n}{N} = \frac{9464*95}{15336} = 59$
Stratum-3	N3=total population of chamokebele	$n_3 = \frac{N3*n}{N} = \frac{2000*95}{15336} = 12$

2.5. The Study Variables

Dependant variable: using of credit loan payment.
Independent variables: sex, age, family size (number of children), educational level, religion, economic status, attitude, number of times borrowed, awareness, marital status.

Table 2. Dependent and Independent variable categories with their coding.

Variables	Categories
Using of credit Loan payment	Yes=0No=1
Age	18-30=031-40=141-50+=2
Sex	Male = 0 Female= 1
Religion	Orthodox=1Muslim=2 Protestant=3Catholic=4 Other=5
Family size (number of children)	One child=0Two Child=1Three Child=2 more than three Child=3
Educational Level	Elementary=0 High School=1 Preparatory=2 Diploma=3Degree=4Master and above=5
Economic Status	Low=0Medium=1High=2
Attitude	Positive=0Negative=1
Awareness	Yes=0No=1
Number Of Times Borrowed	Once=1 Two times=2More than two=0

2.6. Method of Statistical Data Analysis:Logistic Regression Model

Logistic regression is used when the dependent variable is qualitative in nature or categorical. Qualitative response variable are either binary (dichotomous variable) or multiple category. IN this study binary logistic regression was used.

2.7. Binary Logistic Regression

The binary logistic regression model is defined as

$$\theta(x) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k}} \qquad (1)$$

Where β_0 = the constant of the equation
β_i = the coefficient of the ith predictor variable, i = 1,2,…k
X_i=i^{th} explanatory variable of the model.
An alternative form of the logistic regression equation is;

$$logit[\theta(x)] = log\left[\frac{\theta(x)}{1-\theta(x)}\right]$$
$$= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

2.8. Odds Ratio

The odds ratio is the ratio between two odds. The odds of some event happening is defined as the ratio of the probability of occurrence to the probability of non occurrence. That is, the odd of the event E is given by: $odds(E) = \frac{pr(E)}{pr(notE)} = \frac{pr(E)}{1-pr(E)}$

The odds ratio (OR) for two events A and B given by:

$$OR = \frac{odds(A)}{odds(B)} \qquad (2)$$

2.9. Parameter Estimation for Logistic Regression

To estimate the parameters of logistic regression model, the two estimation methods mostly used are maximum likelihood and non-iterative weighted least squares method. When the assumption of normality of the predictors does not hold, the non- iterative weighted least squares method is less efficient. In contrast, the maximum likelihood estimation method is appropriate for estimating the logistic model parameters due to this less restrictive nature of the underlying assumptions.

Consider the logistic regression model $P(x_i) = \dfrac{e^{X_i'\beta}}{1 + e^{X_i'\beta}}$.

$$L(\beta/y) = \prod_{i=1}^{n} P(y_i \mid X_i')$$
$$= \prod_{i=1}^{n}\left[\frac{e^{X_i'\beta}}{1+e^{X_i'\beta}}\right]^{y_i}\left[\frac{1}{1+e^{X_i'\beta}}\right]^{(1-y_i)} \qquad (3)$$

The objective of ML estimation is to get an estimator $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_k)$ of β which maximizes the likelihood function.

3. Result and Discussion

3.1. Descriptive Statistics

Table 3. Frequency distribution of categorical variables.

Variable	Categories	Frequency	Percent
	18-30	40	42.1
Age	31-40	40	42.1
	41-50+	15	15.8
Sex	male	74	77.9
	female	21	22.1
	orthodox	58	61.1
Religion	protestant	33	34.7
	catholic	3	2.1
	other	1	1.1
	elementary	17	17.9
	high school	40	42.1
Educational level	preparatory	13	13.7
	diploma	11	11.6
	degree	13	13.7
	master and above	1	1.1
	single	30	31.6
Marital status	divorced	3	3.2
	Widowed	2	1.1
	married	60	63.2
	no child	33	34.7
	one child	21	22.1
Family size	two child	18	18.9
	three child	12	12.6
	more than three child	11	11.6
Economic status	low	56	58.9
	medium	39	41.1
Attitude	no	69	72.6
	Yes	26	27.4
Awareness	No	69	72.6
	Yes	26	27.4
	once	7	7.4
Number of borrowed	twice	11	11.6
	More than twice	12	12.6

3.2. Inferential Statistics

Table 4. *Summary of personal chi-square test.*

Variable	Category	Using of credit payment		df	Chi-square	Sig.
		No(%)	Yes (%)			
Age	18-30	22(55.0%)	18(45.0%)	2	11.125	0.004
	31-40	33(82.5%)	7(17.5%)			
	41-50+	6(40.0%)	9(60.0%)			
Sex	male	46(62.2%)	28(37.8%)	1	0.611	0.434
	female	15(71.4%)	6(28.6%)			
Religion	orthodox	39(67.2%)	19(32.8%)	4	2.946	0.567
	protestant	20(60.6%)	13(39.4%)			
	catholic	1(50.0%)	1(50.0%)			
	other	1(100%)	0(0.0%)			
Educational level	elementary	5(29.4%)	12(70.6%)	5	21.225	0.001
	high school	22(55.0%)	18(45.0%)			
	preparatory	12(92.3%)	1(7.7%)			
	diploma	10(90.9%)	1(9.1%)			
	degree	11(84.6%)	2(15.4%)			
	master and above	1(100%)	0(0.0%)			
Marital status	single	21(70.0%)	9(30.0%)	4	2.9466	0.564
	divorced	2(66.7%))	1(33.3%)			
	widowed	0(0.0%)	1(100%)			
	married	37(61.7%)	23(38.3%)			
Family size	no child	21(63.6%)	12(36.4%)	4	1.156	0.049
	one child	13(61.9%)	8(38.1%)			
	two child	12(66.7%)	6(33.3%)			
	three child	9(75.0%)	3(25.0%)			
	more than three child	6(54.5%)	5(45.5%)			
Economic status	low	31(55.4%)	25(44.6%)	1	4.653	0.31
	medium	30(76.9%)	9(23.1%)			
Attitude	no	55(79.7%)	14(20.3%)	1	26.356	0.000
	yes	6(23.1%)	20(76.9%)			
Awareness	no	55(79.7%)	14(20.3%)	1	26.356	0.000
	yes	6(23.1%)	20(76.9%)			
Number of borrowed	once	3(42.9%)	4(57.1%)	3	34.028	0.000
	twice	1(9.1%)	10(90.9%)			
	more than twice	3(25.0%)	9(75.0%)			
	no borrowed	54(83.1%)	11(16.9%)			

The hypothesis to be tested under this condition is:

H_o: there are no association between using credit loan payment and explanatory variables.

H_1: Not H_o

Level of significance α=0.05.

Decision rule: If $X^2_{cal} > X^2_{0.05}(1)$, it's reject Ho or null hypothesis and P-value < α = 0.05,it's fail to reject Ho.

For age of respondents, from the above chi-square test table 4, P-value is less than α-level of significance (i.e. P-value 0.004 < α = 0.05) at df = 2 and Chi-square =11.125, it's reject the null hypothesis; we can be concluded that there is association (dependent) between age ofrespondents and using of credit loan.

For sex: since p- value is greater than α- level of significance (i.e. p-value 0.434 >α=0.05) at Chi-square 0.611 with df = 1, then it's fail to reject the null hypothesis at 5% level of significance thus it can be concluded that sex and using of credit loan is no association (independent) between sex of respondent and using of credit loan.

For religion of the respondent: from the Chi-square test above table, since the P-value is greater than common level of significance (i.e P-value = 0.567 >α = 0.05) at value of Chi-square = 2.946 with df = 4, it is fail to reject H_o, we can

be concluded that there isno association between religion of the respondent and using of credit loan.

For educational level of the respondent, since p- value is less than common level of significance (i.e p-value = 0.001< 0.05) at value of Chi-square = 21.225 with df = 5, then it's reject the null hypothesis at 5% level of significance thus it can conclude that the religion of the respondent and utilization of credit loan are dependent or there is association between the religious of the respondent and using of credit loan.

For economic level of the respondent from the above table Chi-square test of hypothesis, P-value is greater thancommon level of significance (i.e. P-value 0.31>α=0.05) itChi-square = 4.653 with df = 1, it is fail toreject the null hypothesis, We can conclude that there is noassociation (independent) between economic level of the respondent and using of credit loan.

3.3. Binary Logistic Regression Analysis

Table 5. *Omnibus Tests of Model Coefficients.*

		Chi-square	df	Sig.
Step 1	Step	83.728	25	.000
	Block	83.728	25	.000
	Model	83.728	25	.000

Omnibus Tests of Model Coefficients gives us a Chi-Square of 83.728 which is significant at 5% level. This is a test of the null hypothesis that adding the predictors to the model has not significantly increased our ability to predict utilization of credit loan. Since our omnibus test is significant we can conclude that adding the predictors to the model has significantly increased our ability to predict utilization of credit loan.

3.4. Model Summary

Table 6. *Cox and Snell Statistic.*

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	40.050(a)	.586	.805

Cox and Snell or Nagelkerke R^2 is an analogous statistic in logistic regression to the coefficient of determination R^2 in linear regression, but not a close analog. The model summary provides some approximations of R^2 statistician logistic regression. Cox and Snell's R^2 attempts to imitate multiple R^2 based on 'likelihood'. In this study Cox and snell R^2 indicate that 58.6% of the variation in the dependent variable using credit loan is explained by explanatory variables.

3.5. Hosmer and Lemeshow Test

Table 7. *Hosmer-Lemeshow test.*

Step	Chi-square	df	Sig.
1	2.843	8	.944

Hosmer-Lemeshow test is used to assess the overall goodness of fit of the fitted model. The P-value is 0.944 for the Hosmer-Lemeshow goodness of fit test is greater than 0.05, we will not reject the null hypothesis that there is no difference between observed and model predicted values, implying that the model estimates are adequate to fit the data at an acceptable level.

Table 8. *Variable in the equation for binary logistic regression.*

		B	S.E.	Wald	df	Sig.	Exp(B)
	Age			9.158	2	0.010	
	Age(1)	-5.551	2.157	6.622	1	0.010	0.004
	Age(2)	-6.900	2.287	9.102	1	0.003	0.001
	Sex(1)	3.508	1.449	5.858	1	0.016	33.380
	Religion			2.895	4	0.576	
	Religion(1)	-12.459	40.992	0.000	1	1.000	0.000
	Religion(2)	-10.787	412.92	0.000	1	1.000	0.000
	Religion(3)	-10.561	92.992	0.000	1	1.000	0.000
	Religion(4)	-32.067	541.459	0.000	1	1.000	0.000
	Education level			11.598	5	0.041	
	Educationlevel(1)	24.354	192.54	0.000	1	1.000	377.260
	Educationlevel(2)	21.421	12.94	4.670	1	0.039	201.463
Step	Educationlevel(3)	13.899	401.95	0.000	1	1.000	1087.527
1(a)	Educationlevel(4)	18.068	492.95	0.000	1	1.000	702.811
	Educationlevel(5)	19.650	402.954	0.000	1	1.000	3542.973
	Maritalstatus			7.183	4	0.127	
	Maritalstatus(1)	-1.468	443.885	0.000	1	1.000	0.230
	Maritalstatus(2)	11.000	40192.885	0.000	1	1.000	5985.188
	Maritalstatus(3)	25.441	56841.383	0.000	1	1.000	111.306
	Maritalstatus(4)	6.503	40192.885	0.000	1	1.000	667.461
	Familysize			8.204	4	0.024	
	Familysize(1)	8.835	3.424	6.657	1	0.010	6871.124
	Familysize(2)	3.938	2.223	3.138	1	0.076	51.317
	Familysize(3)	.374	1.990	0.035	1	0.851	1.454
	Familysize(4)	-1.528	2.406	0.403	1	0.525	0.217
	Economicstatus(1)	0.757	1.489	0.259	1	0.611	2.132
	Attitude(1)	-0.106	5.632	8.534	1	0.046	0.899
	Numberofborrowed			6.430	3	0.092	
	Numberofborrowed(1)	1.344	5.581	0.058	1	0.810	3.833
	Numberofborrowed(2)	9.222	6.352	2.107	1	0.147	1013.365
	Numberofborrowed(3)	3.663	5.794	0.400	1	0.527	38.994
	Constant	-17.360	69616.172	0.000	1	1.000	0.000

Therefore from the above table the p-values of respective variables are less than 0.05 for age of respondents for 31-40 and 41-50+, sex of female respondents, educational level of respondents for high school, family size of the respondents have one child, and attitude of therespondents are significantly association with the response variable. Thus, there is enough evidence to conclude that there is a significant relationship or association at 5% level of significance between predictor variables like age, sex, educational level, family size and attitude of the respondents with credit loan at 5% level of significance.

The fitted logistic regressionmodel is: Logit (pi) =log $(pi/1-pi) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$. So the logit model is used to predict the explanatory variable of the model.

Logit (pi) $= -17.360 - 5.551x_1 + 3.508x_2 + 24.354 x_3 + 8.835 x_4 - 0.106 x_5$

Where x_1=age of the respondent

X_2= sex of respondent is female

X_3 = educational level is high school

X_4=family of one child

X_5=attitude of the respondents

3.6. Interpretation Using Odds Ratio

The odds ratio is equal to 1 when no association exists. It is the odd of success given a certain condition exists divided by the odds of failure given that same condition does not exist. One advantage of the logit function is that it provides an estimate of the odds ratio for each predictor in the model.

- Theestimated odds ratio for age group between 18-30 and 41-50+ is 0.4 and 0.1 respectively as compared with the age group 18-30. This means that using of credit loan on the age interval 31-40 and 41-50+ are greater effect on using credit loan about 0.4 and 0.1 times that of age 18-30 (reference group).
- The estimated odds ratio for sex (1) or female of the respondents is 38.0% as compared with male this means females are higher effect on using credit loan than male.

4. Conclusion and Recommendation

This study considers the factors that affect using of credit loan payment in Arba Minch town Secha sub-city in 2007 E.C. stratified random sampling is used for data collection and analysed bay logistic regression model. From the empirical result credit loan is significantly related with age, sex, education level, family size, awareness, attitude and number of borrow times on the other hand using of credit loan is not significant relation with religion economic status and marital status. Based on the findings of the study the following are recommended

- Governments should expand the awareness about the utilization of credit loan payment for people who live in secha sub-town and also expands the access of insurance company service in their village to minimize the transportation cost and waiting time.
- Governments should give sufficient knowledge about using credit loan in order to people knows the advantage of credit loan and most of the people have the owner of using credit loan.
- Government should give reward for more times borrower in order to facilitate them and people to participate using of credit loan to see those reward people.

- If the urban people get credit on time they can utilize the loan effectively and get a better output. Therefore, timely credit service help people to use the loan in appropriate way and perceived importance of loan by the borrowers and accelerates the chance of the repayment of the loan at proper time.

References

[1] Aryeetey, Ernest, and William Steel, (1995).Saving s collectors and financial intermediation in Ghana. Savings and Development 19(2):189-199.

[2] AnbeseTenaye, (2003).Pattern of credit use and its impact on small farmer's income.A study in Dire Dwa area. Eastern Ethiopia.un published M.Sc.

[3] Beck man, T.N. and Foster.R.S.(1969). Credit and collection, Management.

[4] Befkadu, (2007). Outreach and financial performance of analysis of micro finance institution in Ethiopia. African Economic Conference.

[5] Belay Kebede, (1998). Credit and factor impeding loan payment performance of small holder in central high lands of Ethiopia.

[6] Hussein Hamada, (2007). Farm House holds Economic Behavioral in Imperfect Financial Markets, Doctoral thesis, Swedish University of Agricultural Sciences, Uppsala.

[7] IFC (2012)"Private Sector Boosts Ethiopia's Growth".

[8] KebedeKoomsa, (1995).Agricultural credit Analysis. National Agricultural policy work shop.

[9] Pinkie, (1998). Formal and Informal sector interaction in rural credit markets. Journal of development economics 56:265-280.

[10] Sherraden, M. (2001).Asset building policy and programs for the poor.InT.Shapiro and E. Wolff (Eds.), Asset for the poor.

[11] Solomon Damte,(1996) "credit to poor Households, The case of financing micro enterprises in the DebreBerhan town, North shoa, in the Ethiopian Economy: poverty and poverty alleviation" proceeding of the 5th annual, conference on the Ethiopian economy, Addis Ababa.

[12] TeferiZewdu,(2000) micro finance and the poor: The case of dedebit credit and saving Institution (DESCI).

[13] Vigno, Lawara. (1993). Credit scoring model for development Banks: An African case study, saving and Development vol. XVII.