

Statistical Analysis on the Loan Repayment Efficiency and its Impact on the Borrowers: a Case  
Study of Hawassa city, Ethiopia  
Yonas Shuk, Hawassa University  
E-mail: [yonyshkk@gmail.com](mailto:yonyshkk@gmail.com)  
Department of Statistics, Hawassa University, 05, Hawassa, Ethiopia

---

**Abstract:** The major objective of this study is to assess factors affecting loan repayment efficiency of borrowers and assess impact of efficient utilization of loan for the borrowers in Hawassa city in Ethiopia. Data used for this study was collected through a structured questionnaire. Classical and Bayesian logistic regression technique were used for data analysis. Factor analysis was used to reduce data and to incorporate the major determinants that the efficient utilization of loan have to the borrowers, whereas logistic regression is used to obtained factors affecting loan repayment performance of borrowers and it was extended to the Bayesian frame work using prior information that the parameter follows. Results of the classical binary logistic regression indicate that better repayment efficiency is associated with borrowers: sex, educational status, number of dependent family member, monthly income, loan size, additional source of income, motivation of repayment, time given for repayment, interest rate and screening mechanism when individuals apply for the loan. Also by using Bayesian logistic regression age, loan type, using loan for intended purpose and experience are significant in addition to significant predictors in classical logistic regression. From factor analysis, 27 factor used for impact assessment in which all the factor loaded highly in 7 significant factors like:- Benefit and obstacle related factor, capital effect, saving habit, expenditure, government spending, satisfaction level on the service and consumption change that has been seen after using loan. Thus, in order to improve repayment performance of borrowers, increasing loan size, training and giving some incentive in business areas, increasing awareness in different ways and studying factors which has significant impact on borrowers creditworthiness and giving solution to reduce that problems must be improved.

**Key Words:** *Loan Repayment Efficiency, Loan Impact, SMFI, Logistic Regression, Bayesian Logistic Regression, Multivariate Factor Analysis, Hawassa.*

## **1. Introduction**

Microcredit is the process of lending capital in small amounts to poor people who are traditionally considered unbankable to enable them to invest in self-employment (Kasim and Jayasooria, 2001). The World Bank (2006, p12) describes microcredit as “a process in which poor families borrow certain amounts of money at one time and repay the amount in a stream of small, manageable payments over a realistic time period using social collateral in the short run and institutional credit rule in the long run”.

In Ethiopia where the farming system is at the traditional level and the industrial and service sectors are at their infant stage, the role of microfinance and small enterprises is insignificant in terms of their employment generation capacity, quick production response, adaptation to weak infrastructure, use of local resources and as a means of developing indigenous entrepreneurial and managerial skills for a sustained growth need (Aryeetey, 1994 in Fasika and Daniel, 1997). For small-scale enterprises to grow up to medium and large-scale level, the need for formal credit source is indispensable.

Lack of collateral and the smaller size of the loan demanded by the sector have resulted in a lesser interest on formal financial intermediaries, such as banks to consider it as a potential customer. The higher interest rate charged by some informal money lenders made the financial

problem more unreachable; thus, MFIs were aimed to bridge this gap as their primary objective through MFIs, the poor, especially the informal sector have been proved to be bankable (Ghatak 1998) i.e. they deliver loans to low income peoples through MFIs. However, the recent trend in repayment rate shows deterioration. Its loan recovery rate reduced dramatically from 38% and 64% in 1996/97 to 24% and 31% in 1999/2000 (Michael, 2006) respectively. The default problem mentioned above and the stringent lending criteria used by banks seem paradoxical because, on the one hand only a limited number of borrowers could get credit access and on the other hand a considerable portion of these eligible borrowers are in default problem. Thus, it is good to make an empirical investigation on the factors behind the default problem so that the lending unit could make an appropriate precaution in its lending decision as well as revise its screening criteria.

When the client applies for a loan, then the application can be accepted or rejected by the creditor. The accepted applicant will receive a loan. After a certain period of time, the loan performance and its impact will be assessed as good or bad/efficient or not efficient. The selection mechanism determines whether the application is accepted or rejected and the outcome mechanism determines the loan performance of the accepted application.

In this paper we use different methods to examine factors affecting loan repayment efficiency and the impact of loan for the borrowers. Among many statistical methods that can be used to implement these studies, we use logistic regression to predict the category of outcome for individual cases and to find the best fitting model to describe the relationship between the response and explanatory variables, Bayesian logistic regression is used to predict repayment efficiency of borrowers by including the prior information in the subject and factor analysis is computed to get most significant factor that the impacts of efficient utilization of loan on the borrowers financed by SMFI in Hawassa city and for data reduction.

## **2. Statement of the Problem**

If there is high repayment efficiency, the relationship between the MFI and their client will be good, as Bhattand Tang (2002) argues that high repayment rate helps to obtain the next higher amount of loan and other financial services. Loans taken from credit institutions vary from country to country, region to region, sector to sector. But most credits of developing countries were found to share one common characteristic: Suffer from a considerable amount of default (Kashuliza 1993).

Improving repayment rates helps reduce the dependence of the MFIs on subsidies, which would improve sustainability. It is also argued that high repayment rates reflect the adequacy of MFIs services to client needs (Godquin, 2004). In order to maintain sustainability of MFIs, one important thing is to identify the socio-economic and institutional factors which significantly affect the performance of loan repayment.

There are many socio-economic and institutional factors influencing loan repayment rates in the MFIs. The main factors from the lender side are high-frequency of collections, tight controls, good management of information system, loan officer incentives and good follow ups (Bala, 2011), the size, interest rate charged by the lender and timing of loan disbursement have also an impact on the repayment rates (Oke, et al, 2007). The main factors from the borrower side include socio-economic characteristics such as, gender, educational level, marital status and household income level and peer pressure in group based schemes and etc. SMFI is among the pioneer MFIs in the country providing services in and around the capital city of SNNPRS, which also experiences considerable problem of default.

This study answers the following basic questions:

- I. What are the major socio-economic factors influencing loan repayment efficiency of the borrowers in SMFI?
- II. What are the businesses and loan related factors that influence the repayment performance of the clients?
- III. What are the major problems and challenges faced by the borrowers and lenders in the repayment process in SMFI?

### **3. Objectives of the Study**

The general objective of this study is to examine the efficiency of borrower's loan repayment, dynamic incentives and effects on borrowing decisions and assess factors affecting repayment efficiency of borrowers financed by SMFI in Hawassa city.

#### **The Specific Objectives:**

1. To assess the factors affecting efficiency loan repayment performance of borrowers from different loan products financed by SMFI in the Hawassa City.
2. To identify the factors influencing the repayment of microcredit in Hawassa city from borrowers and lender side.
3. Evaluate the impact of micro credit on household consumption
4. To assess the impact of microcredit on household welfare in regards to income and consumption in Hawassa city.

### **4. Data and Methodology**

#### **4.1 Description of the Study Area and Population**

This study was conducted at Hawassa City, the capital of SNNPR in "Sidama zone" from March 24 to 31, 2012. Hawassa city is one of the administrative city of SNNPR and Sidama Zone which has 8 sub-cities and 40 kebele's. The city is bordered on the south, east and west by the Sidama Zone, on the north by Oromia region. Hawassa city is about 275 kilometers south of Addis Ababa on the paved highway to Kenya through Moyale. According to CSA report (2008) and by using quarter report from economic development and other offices as of (March 2012), the current estimated population size of Hawassa city was 350,000 out of which 180,658 (48.4%) were male and 169,342 (57.6%) were female. The target population consists of all beneficiaries from Sidama micro finance institution in Hawassa city.

#### **4.2 Source and Method of Data Collection**

In this study primary and secondary data was used

- ✓ The secondary data is obtained from the office that is weekly and monthly loan collection modules and some are obtained from operation department of the office.
- ✓ The primary data which is cross-sectional were collected from the target populations by distributing questionnaires on their respective population.

The data was collected by structured household questionnaire that included demographic, social attributes, financial characteristic, service provision variables, household characteristics, main sources of income, assets, credit and saving history, loan utilization, saving and social ties. Open ended questions were included to accommodate unanticipated and broader responses.

#### **4.3 The Study Variable and Description**

**Impact Assessment (IA):** - Is assessing the impact of efficient utilization of loan for the borrowers. In order to study this, we use multivariate factor analysis. The process of IA includes three steps: choosing 'agents' (assessment units), choosing 'outcomes' (assessment indicators) and assessing. Based on this model, we investigate different impact indicator variables and came up with few factor after reduction of data.

**Dependent Variable for Logistic Regression-** selected to be loan repayment efficiency. It is a categorical variable describing efficiency as High, Low or efficient, not efficient.

Efficiency= loan actual repaid / loan to be repaid on time t. (if Efficiency $\geq$ 0.6 High efficiency and low unless)

**Independent Variables are listed as Follows:**

1. **Demographic variables:** age, sex, household size, head of the household, marital status, educational back ground and loan type borrowed
2. **Economic Factors:-** Household assets, income, expenditure, Access to food, health cost, shelter, cost effectiveness, price stability, Information system, market links, turn over, starting capital, current capital, income generated, Amount of saving per month, saving for different purposes.
3. **Loan Utilization and Performance: -** loan type, Availability of other sources of income, repayment period, loan division, purpose of loan, grace period, loan amount, loan reputation, form of disbursement, current loan status.
4. **Institutional and Business:-** Business success, interest, competition, collateral requirement, type of collateral, experience, Social networks, satisfaction level of different services, counseling, loan delivering mechanism, recording, time to repay,
5. **Government Factors:** Taxation, creating job opportunity, legality, Accessing raw material, Supplying place, Financial Support, motivation, screening mechanism, etc.

#### **4.4 Sampling Technique and Sample Size Determination**

A stratified random sampling technique is adopted for this study which involves the division of population into smaller groups, known as strata in such a way that individuals in the same strata are assumed to be homogenous with respect to some characteristics. It is appropriate sampling design for selecting a representative sample, because the borrowers are placed to different types of loan products as strata in the two branches of the city. By considering the loan products as strata, we have found six stratum for this study. The loan product or activities to which SMFI gives loans are agriculture, general loan, housing loans, petty trade, micro and small business, handcraft and services.

#### **Sample Size Determination**

Following (Cochran, 1997), the sample size determination formula adopted for this study is:

$$n = \frac{\sum_{i=1}^k \frac{N_i^2 P(1-P)}{w_i}}{\frac{N^2 e^2}{Z^2} + NP(1-P)} \quad (1)$$

There are different methods of estimating “p”, but for the present study “p” was determined from the results of previous studies. A study which evaluates micro-finance repayment problems in the informal sector in Addis Ababa by Micha'el (2006) has found that the proportion of borrowers with low repayment efficiency is 0.35. This was taken as reference to determining proportion of repayment performance i.e. P (low repayment efficiency) is set to be 0.35.

Having this information, the sample size estimated for this study is:

$$n = \frac{\sum_{k=1}^6 \frac{N_i^2 (p(1-p))}{w_i}}{N^2 / Z^2 e^2 + NP(1-P)} = 316.0076 = 316$$

Finally, 7.5 percent of the sample size i.e.  $23.70 \approx 24$ , was added to compensate none response rate. Thus, the required sample size for this study is 340 beneficiaries which are about 10% of the total population.

Next, the estimated sample size is allocated to each stratum using proportional allocation and the individuals from each loan products are selected by using simple random sampling.

### 5. Multivariate Factor Analysis

The essential purpose of factor analysis is to describe, if possible, the covariance relationships among many variables in terms of a few underlying, but unobservable, random quantities (factors). In this study factor analysis was used to identify the underlying factors or constructs that may influence impact of effective utilization of loan on the borrowers Household welfares.

#### The Orthogonal Factor Model

The factor model postulates that the observable random vector  $X$  with  $P$  components is linearly dependent upon a few unobservable random variables  $F_1, F_2, \dots, F_m$ , called common factors,

and  $p$  additional sources of variation  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$  called errors or specific factors. The factor model is given as:  $X = LF + \varepsilon$

Where  $L_{p \times m}$  is a matrix of unknown constants called factor loading

The coefficient  $l_{ij}$  is the loading of the  $i^{\text{th}}$  variable on the  $j^{\text{th}}$  factor,  $i=1,2,\dots,p$ ;  $j=1,2,\dots,m$ . The unobservable random vectors  $F$  and  $\varepsilon$  satisfy the following conditions.

1.  $F$  and  $\varepsilon$  are independent.

2.  $E(F) = 0$ ,  $Cov(F) = I$ .

3.  $E(\varepsilon) = 0$ ,  $Cov(\varepsilon) = \Psi$ ,

#### Estimation of Loadings

where  $\Psi$  is a diagonal matrix.

4.  $Cov(\varepsilon, F) = 0$

There are two most popular methods of parameter estimation in multivariate analysis, the PC method and the maximum likelihood method. The solution from either method can be rotated in order to simplify interpretation of factors. However, for this study, we consider the principal component method.

#### The Principal Component Method

The spectral decomposition of covariance  $\Sigma$  having eigenvalue-eigenvector pairs  $(\lambda_i, e_i)$  with  $\lambda_1 > \dots > \lambda_m > 0$  is given as  $\Sigma = \lambda_1 e_1 e_1^T + \lambda_2 e_2 e_2^T + \dots + \lambda_p e_p e_p^T$

From above equation, we can obtain the loading  $L = (\sqrt{\lambda_1} e_1, \sqrt{\lambda_2} e_2, \dots, \sqrt{\lambda_p} e_p)$

In applying the principal component to perform factor analysis, we use the sample covariance matrix  $S$  of the sample correlation matrix. observe that  $s_{11} + s_{12} + \dots + s_{pp} = tr(s) = \text{trace of sample covariance matrix}$  and  $\hat{\lambda}_1 + \hat{\lambda}_2 + \dots + \hat{\lambda}_p = p = \text{trace of sample correlation matrix}$ , where,  $\hat{\lambda}_i$ 's,  $i=1, \dots, p$  are the estimated eigenvalues of  $S$ .

$$\left[ \begin{array}{l} \text{The proportion of total sample} \\ \text{variance due to } j^{\text{th}} \text{ factor} \end{array} \right] = \frac{\hat{\lambda}_j}{tr(s)}, \text{ for factor analysis of sample covariance}$$

$$\left[ \begin{array}{l} \text{The proportion of total sample} \\ \text{variance due to } j^{\text{th}} \text{ factor} \end{array} \right] = \frac{\hat{\lambda}_j}{p}, \text{ for factor analysis of correlation}$$

Finally, we choose factors having eigenvalues larger than one.

## 6. Multiple Logistic Regression Model

Logistic regression analysis extends the techniques of multiple regression analysis to research situations in which the outcome variable is categorical. Here, the response variable is binary, such as (efficient or not efficient).

Consider a collection of  $k$  independent variables denoted by a vector  $\mathbf{X}=(X_1, X_2, \dots, X_k)$ .

Let the conditional probability that the outcome of interest in a study is “success” be denoted by  $P(Y=1/\mathbf{X}=\mathbf{x})=P(\mathbf{x})$ .

The ratio of the probability of: success ( $Y_i=1$ )→  $P(\mathbf{x}_i)$  to that of failure ( $Y=0$ )→ $1-P(\mathbf{x}_i)$  is given by:  $\frac{P(x_i)}{1-P(x_i)}$  is known as the odds of a success.

In terms of the odds, the logistic model can be written as:

$$\frac{P(x_i)}{1-P(x_i)} = \exp(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}), \quad i = 1, 2, \dots, n$$

Which means that  $\exp(\beta_j)$ ,  $j=1, 2, \dots, k$  the probability of belonging to one group or event occurring divided by the probability of not belonging to that group or is the factor by which the odds of occurrence of a success change by a level change in the  $j^{\text{th}}$  independent variable.

In which case,

$$P(X_i) = \frac{e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}}}, \quad i = 1, 2, \dots, n$$

It is obvious that the response variable and the predictors are not linearly related. However, to have a linear relationship we can use the logit transformation.

Thus, the transformation of the logistic regression is the logit transformation of  $P(x_i)$ , and is given as:

$$\begin{aligned} \text{logit}(X_i) &= \log \left[ \frac{P(X_i)}{1-P(X_i)} \right] = \log(e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}}) \\ &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}, \quad i = 1, 2, \dots, n \end{aligned}$$

Fitting the model requires the estimates of the values of parameters  $\beta=(\beta_0, \beta_1, \beta_2, \dots, \beta_p)^t$ .

We estimate the parameters using maximum likelihood estimation method.

$$L(\beta) = \prod_{i=1}^n \theta(x_i) = \prod_{i=1}^n [P_i^{y_i} (1 - P_i)^{1-y_i}] = \prod_{i=1}^n \left[ \frac{e^{x^t \beta}}{1 + e^{x^t \beta}} \right]^{y_i} \left[ \frac{1}{1 + e^{x^t \beta}} \right]^{(1-y_i)} \quad \text{Where: } \mathbf{x}^t = (1, x_{i1}, x_{i2}, \dots, x_{ik}), \quad i=1, 2, \dots, n$$

## 7. Bayesian Logistic Regression

Bayesian logistic regression extends logistic regression in to a Bayesian framework (Xu and Akella 2008). Bayesian inference, which allows ready incorporation of prior beliefs and the combination of such beliefs with statistical data, is well suited for representing the uncertainties in the value of explanatory variables (Jaakkola and Jordan 1996).

Mathematically, the conditional probability of observed data  $D$  given parameters  $\beta$  relates to the converse conditional probability of parameters  $\beta$  given observed data  $D$ :

$$P(\beta/D) = \frac{P(\beta, D)}{P(D)} = \frac{P(D/\beta)P(\beta)}{P(D)}$$

Where:-  $P(\beta, D)$  is a joint probability distribution for  $\beta$  and observed data  $D$ ;  $p(\beta)$  is a prior probability for  $\beta$ ,  $P(\beta|D)$  is a posterior probability for parameters  $\beta$ ;  $P(D|\beta)$  is the likelihood function, and  $P(D)$  is the probability distribution of observed data  $D$ .

In Bayesian framework, there are three key components associated with parameter  $\beta$ : the prior distribution, the likelihood function, and the posterior distribution. These three components are formally combined by Bayes rule:  $f(\beta/y) \propto \text{Likelihood} \times \text{Prior}$

### Likelihood Function

Let  $y_1, y_2, \dots, y_n$  be independent Bernoulli trials with success probabilities  $P_1, P_2, \dots, P_n$ , that is  $y_i = 1$  with probability  $P_i$  or  $y_i=0$  with probability  $1- P_i$ , for  $i= 1, 2, \dots, n$ . The trials are independent, the joint distribution of  $y_1, y_2, \dots, y_n$  is the product of  $n$  Bernoulli probabilities given as:

$$f(y/\beta) = \prod_{i=1}^n [P_i^{y_i} (1 - P_i)^{1-y_i}]$$

Where,  $p_i$  represents the probability of the event for subject  $i$  who has covariate vector  $X_i$ ,  $y_i$  indicates the presence,  $y_i=1$ , or absence  $y_i=0$  of the event for that subject.

$$P_i = \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}}$$

where:  $P_i$  = the probability of  $i^{\text{th}}$  employees being low efficient, since individual subjects are assumed independent from each other likelihoods function over a given data set of subjects is:

$$f\left(\frac{y}{\beta}\right) = \prod_{i=1}^n \left\{ \left[ \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}} \right]^{y_i} * \left[ 1 - \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}} \right]^{(1-y_i)} \right\}$$

### Prior Distribution

For this study, we use the most common priors for logistic regression parameters, which is a normal distribution with mean  $\mu$  and covariance matrix  $\Sigma$ . That is;  $f(\beta) \sim N(\mu, \Sigma)$ .

The most common choice for  $\mu$  is zero vectors, and  $\Sigma$  is usually chosen to be a diagonal matrix ( $\Sigma = \text{diag}(\sigma_0^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2)$ ) with large variances that to be considered as non-informative prior, common choices for the variances ( $\sigma_j^2$ ) is in the range from 10 to 100.  $f(\beta_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{1}{2}\left(\frac{\beta_j - \mu_j}{\sigma_j}\right)^2\right)$

### The Posterior Distribution

Given the likelihood and the prior distribution given above, the posterior distribution of the Bayesian logistic regression contains all the available knowledge about the parameters in the model like:

$$f(\beta/y) \propto \prod \left[ \frac{P_i^{y_i} (1 - P_i)^{(1-y_i)} P(\beta)}{P(X_1, X_2, \dots, X_p)} \right] \prod_{i=1}^p [P_i^{y_i} (1 - P_i)^{(1-y_i)} P(\beta)]$$

$$= \prod_{i=1}^K \left\{ \left[ \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}} \right]^{y_i} \left[ 1 - \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}}} \right]^{(1-y_i)} \right\} \times \prod_{i=0}^K \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{1}{2}\left(\frac{\beta_i - \mu_i}{\sigma}\right)^2\right)$$

Where  $f(\beta|y)$  are the posterior distribution which is the product of likelihood and the normal prior distributions for the  $\beta$  parameters of the logistic regression.

Estimation of  $\beta$  on the posterior distribution may be difficult, for this reason we need to use non-analytic method. The most popular method of simulation technique is Markov Chain Monte Carlo (MCMC) methods.

### Markov Chain Monte Carlo (MCMC)

Simulation is a general computational method in Bayesian inference to obtain a sequence of random samples from a probability distribution. This method is based on drawing values of parameters  $\beta$  from approximate distributions, and then correcting those draws to better approximate the target posterior distribution,  $P(\beta|D)$ .

Standard Monte Carlo methods produce a set of independent simulated values according to some desired probability distribution.

Markov chain is a stochastic process with the property that any specified state in the series,  $\beta^{[t]}$  is dependent only on the previous value of the chain,  $\beta^{[t-1]}$  and is therefore conditionally independent on all other previous values. This can be stated more formally as:

$$P(\beta \in A | \beta^0, \beta^{[1]}, \dots, \beta^{[t-2]}, \beta^{[t-1]}) = P(\beta^{[t]} \in A | \beta^{[t-1]})$$

The advantage of this notation is that it subsumes both the continuous state space as well as discrete state space. When the state space is discrete, then  $K$  is a matrix mapping,  $k \times k$  for “k” discrete elements in  $A$ , where each cell defines the probability of a state transition from the first term to all possible states:

$$P_A = \begin{bmatrix} P(\beta_1, \beta_1) & \dots & P(\beta_1, \beta_k) \\ \vdots & \ddots & \vdots \\ P(\beta_k, \beta_1) & \dots & P(\beta_k, \beta_k) \end{bmatrix}$$

Where the row indicates the chain is at this period and the column indicates where the chain is going in the next period. The rows of  $P_A$  sum to one.

The Chapman-Kolmogorov equations specify how successive events are bound probabilistically. These are given here for both discrete and continuous state spaces:

$$P^{(m_1+m_2)}(x, y) = \sum_{\text{all } z} P^{m_1}(x, y) P^{m_2}(z, y) \\ - \text{discrete case}$$

$$P^{(m_1+m_2)}(x, y) = \int_{\text{range } z}^{\infty} P^{m_1}(x, y) P^{m_2}(x, y) dz - \text{continuous case}$$

The above equation can be represented as a series of segmented matrix multiplications:

$$P^{l(m_1+m_2)} = P^{m_1} P^{m_2} = P^{m_1} P^{m_2-1} P^{m_2}$$

The final basic notational characteristic of Markov Chains that we will provide here is the marginal distribution at some step  $m^{\text{th}}$  from the transition kernel. For the discrete case, the



marginal distribution of the chain at the “m” step is obtained by inserting the current value of the chain,  $\beta_i^{(m)}$  in to the row of the transition kernel for the m<sup>th</sup> step,  $p^m$ :

$$\pi(\beta)=[P^m(\beta_1), P^m(\beta_2), \dots, P^m(\beta_k)]$$

So the marginal distribution at the first step of the Markov chain is given by:

$$\pi^1(\beta) = \pi^0(\beta)P_1$$

Where  $\pi^0$  is the initial starting value assigned to the chain and  $p^1 = p$  is the simple transition matrix.

A neat consequence of the defining characteristic of the transition matrix is the relationship between the marginal distribution at some (possibly distant) step and starting value:  $\pi^n = P\pi^{n-1} = P(P\pi^{n-2}) = \dots = P^n\pi^0$

Since it is clear here that successive products of probabilities quickly result in lower probability values, the property above shows how Markov chains eventually “forget” their starting points. The marginal distribution for the continuous case is only slightly more involved since we cannot just list as a vector quantity:

$$\pi^m \beta_j = \int_{\theta} P(\beta, \beta_j) \pi^{m-1}(\beta) d\beta$$

This is the marginal distribution of the chain, currently on point  $\beta_j$  at step m.

### Stationary Distribution

Define  $\pi(\beta)$  as the stationary distribution of the Markov chain for  $\beta$  on the state space A. We denote  $P(\beta_i, \beta_j)$  the probability that the chain will move from  $\beta_i$  to  $\beta_j$  at some arbitrary step t from the transition kernel, and  $\pi^t(\beta)$  as the marginal distribution. Thus, the stationary distribution is a distribution satisfying:

$$\sum_{\beta_i} \pi(\beta_i) P(\beta_i, \beta_j) = \begin{cases} \pi^{t+1}(\beta_j); & \text{discrete case} \\ \int \pi(\beta_i) P(\beta_i, \beta_j) d\beta; & \text{in continuous case} \end{cases}$$

Once the chain reaches its stationary distribution, it stays and moves around, or “mixes” throughout the subspace according to marginal distribution, forever. Then all we need to do is let it wander about this subspace for a while, producing empirical samples to be summarized. The most commonly used MCMC techniques are Metropolis-Hasting and Gibbs sampler techniques.

### The Gibbs Sampler Algorithm

The Gibbs sampler (David, 2006) is the most widely used MCMC technique. It is a transition kernel created by a series of full conditional distributions that is a Markovian updating scheme based on conditional probability statements. The set of full conditional distributions for  $\beta$  are denoted  $\pi(\beta)$  and defined by  $\pi(\beta) = \pi(\beta | \beta_i)$  for  $i = 1, 2, \dots, k$ , where the notation  $\beta_i$  indicates a specific parametric form from  $\beta$  without the  $\beta_i$  coefficient. These requirements facilitate the iterative nature of the Gibbs sampling algorithm described as:

I. Start with an initial value:  $\beta^0 = [\beta_0^{(0)}, \beta_1^{(0)}, \beta_2^{(0)}, \dots, \beta_k^{(0)}]$

II. Sample for each  $i = 0, 1, \dots, n-1$ : Generate  $\beta_0^{(i+1)}$  from

$$f(\beta_0 | \beta_1^{(i)}, \beta_2^{(i)}, \beta_3^{(i)}, \dots, \beta_k^{(i)})$$

Generate  $\beta_1^{(i+1)}$  from

$$f(\beta_1 | \beta_0^{(i+1)}, \beta_2^{(i)}, \beta_3^{(i)}, \dots, \beta_k^{(i)}) :$$

Generate  $\beta_k^{(i+1)}$  from

$$f(\beta_k | \beta_0^{(i+1)}, \beta_1^{(i+1)}, \beta_2^{(i+1)}, \dots, \beta_{k-1}^{(i+1)})$$

III. Return  $\beta^{(1)}, \dots, \beta^{(n)}$

In this study, we used Win BUGS software to approximate the marginal posterior distributions for each parameter.

## **8. Results and Discussion**

Out of 340 borrowers considered in the analysis, 38.53% beneficiaries are efficient on repayment and 61.47% are not efficient at the time of data collection (Appendix A1, Table 4.1). Of the total sample, 11.8% of the clients borrowed for Agricultural Products, 22.6% of clients for petty trades, 21% for Micro and small enterprises, 8.8% for Hand craft and Service, 10.9% for General Loan products and 23.8% for Housing products. With regard to the sex composition, 38.2% were female and 61.8% were male borrowers.

### **8.1 Results of Factor Analysis on Impact Loan Attributes to the Borrowers**

Before factor analysis is conducted, the reliabilities of the variables were checked against the recommended standards (Cronbach  $\alpha \geq 0.70$ ) mainly to ensure that they are reliable for the factor analysis (Nunnally, 1967). Factor analysis using principal components has been applied using 27 efficient lone utilization impact factors that were obtained from the household survey. Orthogonal factors were obtained using varimax rotation. Only those factors with Eigen value greater than 1.0 and high cronbach  $\alpha$  coefficients are considered.

1<sup>st</sup> factor: This mostly shows high loading on the experience and peer related factor which was obtained from survey data and can be labeled as Benefit and obstacle related impact of loan for the borrowers. All the factors include number of reputation (number of times loan was received, age of the borrowers, peer effect and distance of the company/Home from the institution and it can be said together as maturity on the loan utilization or Household improvement factors.

2<sup>nd</sup> factor consists of, income, loan size, loan type and type of collateral used as guarantee. Thus it is labeled as income (capital) dimension of microcredit loan impact on the borrowers in Hawassa city.

3<sup>rd</sup> factor shows high loading on saving related factor like, save for another personal cases, save to get another loan, save to strengthen the business, save for investment, save since obligation of the institution, save for insurance (death, health care, accident) and we can label as saving dimension of loan impact.

4<sup>th</sup> factor includes screening mechanism, counseling service, motivation and support which can be labeled as government role on loan impact.

5<sup>th</sup> factor: Includes level of satisfaction on the different service given for the clients in the organization like time scheduling for repayment, interest rate of the organization, handling ways of the customers, satisfaction level of inflation and the like which can be labeled as satisfaction level of the customers.

6<sup>th</sup> factor: Expenditure related factors, which was obtained from survey data and can be labeled as Expenditure cost on microfinance loan impact.

7<sup>th</sup> factor: Includes some of changes/impacts of different consumption and cost effectives cases of loan impact; like business change after using loan, Improvement in food consumption, improvement in health care, improvement in different facilities, increasing size and quality of trade and the like which can be labeled as consumption dimension of loan impact on the borrowers.



Table 1. Results of Principal Component Factor Analysis of Items related to Efficient Utilization of Loan Impact (Cronbach's  $\alpha = 0.712$ )

Common Factors: Components							
Accounted for 75.824%	F1	F2	F3	F4	F5	F6	F7
Eigen Value	4.494	3.185	2.389	1.749	1.419	1.383	1.28
Original Variables Having Commonalities			>0.50				
% Variance Explained	24.885	18.598	13.165	6.215	5.5	4.405	3.056
Number Repetition	0.8078						
Age	0.7874						
Experience	0.6996						
Distance from home	0.6348						
Peer effect	0.7894						
Mincom		0.8484					
Type collateral		-0.623					
Amount		0.723					
Loan Type		0.4552					
Pricing		-0.528					
Save for investment			0.7263				
Save to get another loan			0.6587				
Save for insurance purpose			0.7399				
Save to strengthen business			0.6347				
Ways Selecting Applicants				-0.805			
Training and counseling				0.7277			
Motivation and support				0.7841			
Expenditure for food					0.798		
Expenditure for housing					0.545		
Expenditure for consumption					0.5061		
Satisfaction level by timing						0.6106	
Satisfaction by service						0.6908	
Satisfaction level on inflation						0.4474	
Improvement of business							0.6103
Food consumption							-0.482
Health care system							0.4803
Additional facilities							0.436

**8.2 Determinants of loan Repayment efficiency using Logistic Regression** The significant predictors of repayment efficiency of the borrowers using forward likelihood ratio method for variable selection in

multiple logistic regression models were:- Sex of borrowers, Family size, Educational status, Amount of loan that they have borrowed, Tax laid by the government from different direction, Motivation of the repayment, Monthly income, Time or duration given to repay loan, Presence of additional income, Interest rate laid by the institution, Source of additional income, and Screening mechanism when the borrowers apply for the loan.

Since most significant predictors are categorical, the values of the Wald statistics and the odds ratios for each category with their respective probabilities are given in Table 2 below. Here to interpret the odds ratio, we use last category as a reference group.

From the result, since the probability of Wald statistics for each of the above 12 covariates was less than the level of significance 0.05, we mainly focus on the categories of these variables to interpret the effects of each covariate using the estimated odds ratio.

The result shows that repayment efficiency of borrowers is associated with sex of the clients, since  $p$ -value=0.029 and odds ratio was 0.41. This indicates that, females are (1-0.405) which is 0.5905 times less efficient on repayment than male borrowers. This may be due to inefficiency of female borrowers to actively participate in the business activities in comparison with male, inactive participation of females in different areas, low educational status of females, culture etc.

Family size also has significant contribution for repayment efficiency of borrowers ( $p=0.01$ ). In household wise, small family size ( $\leq 3$ ) are 6.31 times more likely to repay efficiently than those with more than 10 family members. Also those having family size from (4-6) are 6.08 times more likely to repay efficiently than those having above 10 dependent family members. Regarding educational status, borrowers with 2<sup>nd</sup> cycle of elementary school are (1-0.51) times less likely efficient on repayment than the reference category (degree and above), those in high school, certificate

and diploma are  $(1-0.13) = 0.87$ , 0.91 and 0.73 times less likely efficient on repayment than the reference categories. The positive sign for the logit coefficient of the covariate indicates that as educational status of the borrower increases the repayment efficiency also increases. Similarly, the logit coefficient for illiterate and less educated clients is negative, indicating that low repayment efficiency is associated with low educational status..

Monthly income is also significant factor among 24 predictors which are used to compute multiple logistic regressions. From different categories, individuals whose average monthly income lower are not efficient on repayment, where as those whose monthly income is (801-1200) and (1200-1500) with OR=0.29, 0.24 respectively were efficient on repayment even if they have less effect in comparison to reference categories ( $>1700$  birr).

When we came to presence of additional income, those having additional income are 47.134 times more likely to pay back with better efficiency than those who have only one income source. Thus, it is good to divert source of income in different direction. If a husband job is government employee, and then may be his wife be merchant, or technician or they can have additional work that they can run jointly.

Also, the amount of loan individuals have borrowed, tax laid by the government, motivating the borrowers in different means, duration given to repay the loan back, clear and fair screening mechanism and to smaller interest rate has significant contribution on the repayment efficiency of borrowers since their  $p$ -values are less than 5% at 5% level of significance.

Table 2: Results of the Final Multiple Logistic Regression Model

Parameter	$\hat{\beta}$	Std. Error	Wald $\chi^2$	Df	Sig.	Exp( $\hat{\beta}$ )	95% Confidence for Exp(B)	
							Lower	Upper
(Intercept)	1.991	0.5929	11.27664	1	0.011	7.33	5.05	9.607
[Sex=Female]	-0.903	0.414	4.76	1	0.029	0.405	0.18	0.913
[Sex=Male]	Ref							
[Fams=1]	1.841	0.647	8.1	1	0.004	6.31	1.77	22.41
[Fams=2]	1.92	0.6616	8.42	1	0.004	6.82	1.87	24.94
[Fams=3]	0.806	0.621	1.68	1	0.194	2.24	0.66	7.56
[Fams=4]	Ref							
[EduSta=1]	-0.677	0.199	0.32	1	0.573	0.51	0.05	5.33
[EduSta=2]	-0.734	0.2057	0.37	1	0.542	0.48	0.05	5.10
[EduSta=3]	-0.911	0.229	15.83	1	0.003	0.4	0.05	0.98
[EduSta=4]	-2.005	0.9005	4.96	1	0.026	0.13	0.02	0.79
[EduSta=5]	-2.431	0.976	6.2	1	0.013	0.09	0.01	0.60
[EduSta=6]	-1.32	0.9348	1.99	1	0.036	0.27	0.04	0.99
[EduSta=7]	Ref							
[Mincom=1]	-0.58	0.898	2.12	1	0.599	0.56	0	1.975
[Mincom=2]	-1.494	0.8397	3.17	1	0.019	0.22	0.04	0.88
[Mincom=3]	-1.228	0.6068	4.09	1	0.043	0.29	0.09	0.96
[Mincom=4]	-1.436	0.4973	8.34	1	0.004	0.24	0.09	0.63
[Mincom=5]	Ref							
[Adin=1]	3.853	0.757	25.906	1	0.011	47.134	31.05	53.97
[Adin=0]	Ref							
[SAI=1]	2.152	0.9251	5.41	1	0.02	8.6	1.4	52.73
[SAI=2]	3.107	0.7412	17.572	1	0	22.354	6.49	56.07
[SAI=3]	2.047	0.984	4.33	1	0.037	7.75	1.13	53.31
[SAI=4]	2.073	0.942	4.84	1	0.028	7.95	1.25	50.38
[SAI=5]	1.576	0.921	2.925	1	0.194	4.84	0.45	52.27
[SAI=6]	Ref							
[Amount=1]	-2.811	0.2662	111.508	1	0.026	0.06	0.01	0.72
[Amount=2]	-2.282	0.8072	7.99	1	0.005	0.1	0.02	0.50
[Amount=3]	-2.703	0.6589	16.8288	1	0.029	0.067	0.004	0.412
[Amount=4]	-2.26	0.8372	7.28	1	0.007	0.1	0.02	0.54
[Amount=5]	-0.251	0.0609	16.987	1	0.003	0.78	0.24	0.97
[Amount=6]	Ref							
[Tax=0]	-1.656	0.4234	15.301	1	0	0.191	0.083	0.438
[Tax=1]	Ref							
[Mrepay=1]	-2.595	0.5002	26.914	1	0.002	0.0746	0	0.19
[Mrepay=2]	0.433	0.3316	0.106	1	0.745	1.54	0.11	20.98
[Mrepay=3]	-0.434	0.5072	0.731	1	0.39	0.65	0.24	1.75
[Mrepay=4]	-1.22	0.5172	5.56	1	0.018	0.03	0.11	0.81
[Mrepay=5]	-1.972	0.845	5.445	1	0.02	0.14	0.03	0.73
[Mrepay=6]	Ref							
[Trepay=0]	1.186	0.4534	6.844	1	0.009	3.27	1.35	7.96
[Trepay=1]	Ref							
[INt=1]	1.08	0.5987	3.256	1	0.012	2.95	1.91	9.52
[INt=2]	0.854	0.4245	4.0472	1	0.025	2.349	1.07	3.73
[INt=3]	Ref							
[ScrM=1]	-0.771	0.292	6.97177	1	0.029	0.05	0.11	0.93
[ScrM=2]	-1.801	0.742	5.89	1	0.015	0.02	0.04	0.71
[ScrM=3]	ref							

Dependent Variable: Loan Repayment efficiency, df=degrees of freedom, Sd= standard error

### 8.3 Determinants of loan Repayment efficiency using Bayesian Logistic Regression

The Bayesian model used here is normal-normal, in which the coefficients are assumed to follow a normal distribution with normal distributed uninformative priors, we assume that the regression parameters of interest all follow a normal distribution with mean = 0 and precision = 1.0e-3 and the inverse Gamma distribution as a prior for  $\sigma^2$  with shape parameter 0.01 for coefficient parameters including constant terms in the model. Since, in Bayesian estimation, the variance of the prior distribution has a great effect in the accuracy of the estimates, we have used uninformative priors to compare the models with different prior variances using DIC value. We apply here three different prior variances. In general, using the model specification Tool, 3 parameter chains with different initial values were set up to be sampled for 40,000 iterations each. The first 20,000 iterations were discarded from each chain (as Burn in since the data converged around 20,000 iterations), leaving a sample of around 70608 to summarize the posterior distribution. In order to minimize autocorrelation, we use every third (thin=3) sample after convergence as it was shown in plot below.

Fig. 2. Time Series Plot

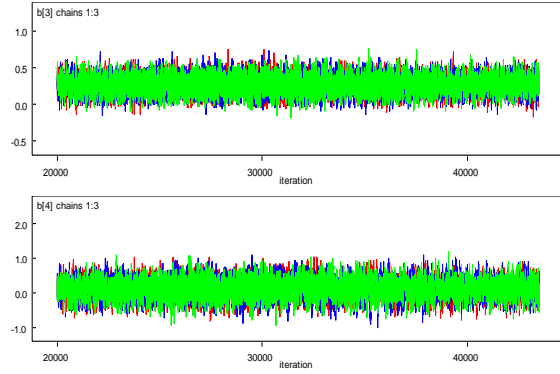
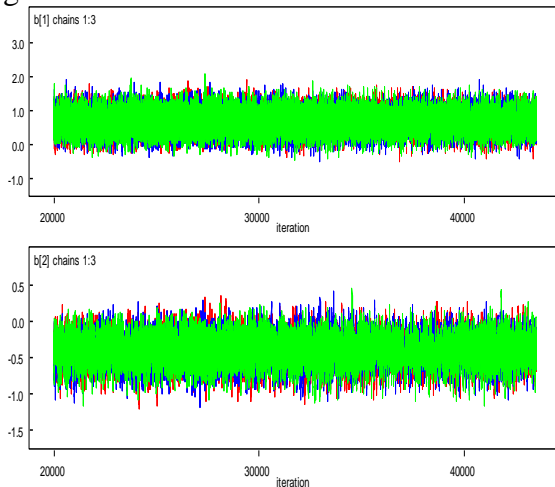


Fig. 3. Kernel Density

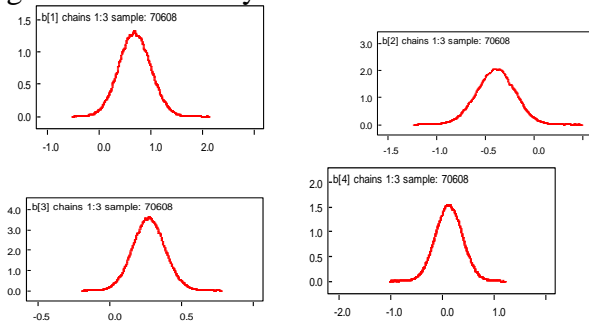


Fig 4. Gelman Rubin Statistic

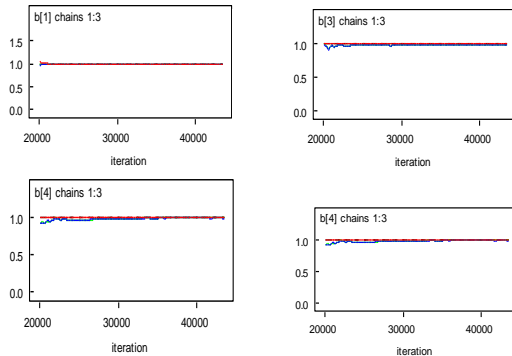
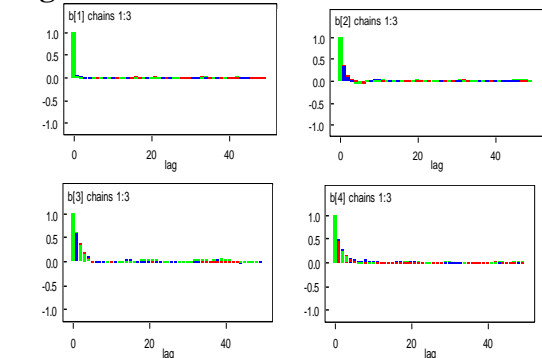


Fig 5. Autocorrelation Plots



From the results of posterior summaries of Bayesian logistic regression model, constant (alpha), the coefficient for sex, age, family size, educational status, monthly income of

borrowers, presence or absence of additional income, source of additional income, amount of loan the beneficiaries have borrowed, tax laid by the government, interest rate the borrowers will pay for the credit, motivating repayment by government, time given to repay loan, loan type, using loan for intended purpose, experience and screening mechanism when borrower apply for the credit are significant efficiency factors for the outcome variables (loan repayment efficiency).

Furthermore, the negative sign of the posterior mean implies that the risk for low repayment was less in comparison to variables having positive coefficient, since the exponents of negative value will be small number which is less than one but not negative and those covariates having positive mean have a higher effect on the repayment efficiency.

When we come to each significant predictors; - sex of borrowers is significant which indicates that being male borrower is more likely to become efficient than being female borrowers, since credible interval of coefficient beta ( $\beta$ ) does not contain zero. Age is also significant predictor of repayment efficiency since OR = 0.6720, Thus, those with lower age categories are more likely to be efficient in repayment than elders. This is because of the younger groups are more actively participate in different business and also has many chance to be involved in different works simultaneously and they have no many dependents i.e. potential youth's who are below 45 years are more active in repayment. In regards to experience, those who have many experiences on the credit are efficient on repayment than those who has only one year experience (reference category).

Actually, as we have seen very small Monte Carlo (MC) error (less than 5% times the posterior standard deviation for all logit coefficients of the explanatory variables) indicates the good model fit (good estimate of the posterior mean and standard deviation). Thus, the model was good fitted model and

good convergence was attained as we have seen in four plots.



Table . Summary Statistics for Bayesian Logistic Regression

Explanatory Var	Node	Mean	$Exp(\beta)$	Sd	Sd*5%	MC e	2.50%	97.50%	Sample
Constant	Alpha	3.102	22.24	2.34	0.117	0.081	1.449	7.633	70608
Sex**	b[1]	0.285	1.329	0.31	0.01546	0.002	0.105	0.893	70608
Age*	b[2]	-0.4	0.672	0.2	0.00978	0.002	-0.79	-0.02	70608
Edu. Status**	b[3]	0.368	1.128	0.11	0.00537	0.002	0.258	1.163	70608
Marital Status	b[4]	0.121	1.128	0.26	0.01293	0.004	-0.38	0.627	70608
Family Size**	b[5]	-0.9	0.407	0.41	0.02032	0.012	-1.71	-0.12	70608
Loan Type*	b[6]	0.347	1.414	0.1	0.0048	0.001	0.162	0.538	70608
Monthly Inc**	b[7]	0.273	1.314	0.16	0.00782	0.002	0.032	0.979	70608
Add. Income**	b[8]	1.273	3.572	0.47	0.02343	0.005	0.353	2.195	70608
S.A A Income**	b[9]	0.026	1.026	0.1	0.0052	0.001	0.176	0.229	70608
Loan size **	b[10]	0.271	1.311	0.11	0.0056	0.001	0.052	0.494	70608
Inflation	b[11]	-1.47	0.231	0.57	0.02826	0.013	-2.61	-0.385	70608
Tax**	b[12]	-0.34	0.71	0.29	0.01452	0.002	-0.92	-0.121	70608
job Satisfaction	b[13]	0.373	1.452	0.37	0.01842	0.003	-0.35	1.096	70608
Number repet	b[14]	-0.56	0.57	0.38	0.01898	0.01	-1.32	0.183	70608
Motivation rep**	b[15]	0.384	1.468	0.11	0.00527	0.001	0.2	0.761	70608
Time repay**	b[16]	0.428	1.534	0.3	0.01501	0.003	0.167	1.355	70608
Interest**	b[17]	-0.23	0.796	0.21	0.01049	0.003	-0.78	-0.165	70608
SCM**	b[18]	-0.79	0.456	0.25	0.01248	0.002	-1.28	-0.305	70608
LIP	b[19]	-0.71	0.494	0.31	0.01559	0.003	-1.33	0.101	70608
Purpose of loan	b[20]	0.21	1.234	0.09	0.00472	0.001	-0.03	0.399	70608
Competition	b[21]	0.118	1.125	0.09	0.00446	0	-0.03	0.321	70608
Gov. incent	b[22]	-0.15	0.858	0.11	0.00531	0.001	-0.36	0.052	70608
Loan Inadequacy*	b[23]	0.685	1.983	0.31	0.01528	0.003	0.092	1.288	70608
Experience*	b[24]	0.232	1.261	0.15	0.00729	0.001	0.054	0.518	70608

Significant in Bayesian logistic regression only (\*) and (\*\*) represents Significant in both Bayesian and classical logistic regression

## 9. Discussions

This study applies factor analysis, classical and Bayesian logistic regression approach to make inference and draw conclusion based on the

data on hand and the prior information that the parameter follows.

According to the results, about 38.5% of the respondents were not efficient at the time of data collection. Out of the beneficiaries who

were not efficient at the time of data collection, 55.5% and 30% were females and elders (age above 55 years) respectively.

The paper also tries to identify impact of efficient utilization of loan on the borrowers by using PCFA. The 27 variables representing factors of loan efficient utilization impact on the beneficiaries are reduced to 7 factors following the factor analysis.

These factors are: Benefit and obstacle related factors which accounts about 25% of total variation consist of factors like repetition, experience in the business, distance from work place to institution and peer effect which indicates the efficient loan utilization impact of the borrowers. Similarly, Diamond (1991) argues experience, reputation, peer effect and age are the impact of experience on loan efficiency and he named as past experience related factor and also, Sahile (2007) identified internal and external factors as factors of loan impact. The second factor accounted about 19% of total variation and mostly consists of economic factors like income, loan size, type of collateral, loan type and pricing and can be labeled as capital factor which efficient utilization of loan has for the borrowers. Third factor which accounts about 13% total variation consists mostly saving for different purposes and labeled as saving impact score. Government impact score, Expenditure impact score, satisfaction level of service impact score and consumption change in social and economic aspects of life impact score are the seven identified factors of impact of efficient utilization of loan for the borrowers. Similarly, Bala (2011), identified seven main factors from 27 items in which staff coordination and customer target are highly dominant impact of loan. Mohammad and Sarker (2009) identified seven main factors from empirical review of microcredit program in Bangladesh from 26 factors by using PCFA. In general, past experience and obstacle, good saving habit,

high capital amount, satisfaction on the service, government role, and change in consumption level after using loan efficiently by decreasing expenditure cost has positive impact that can be seen from efficient utilization of loan by the borrowers.

The most important covariates identified in the multiple logistic regressions are sex, family size, educational status, monthly income, loan size, additional income, source of additional income, tax, motivation of repayment, time to repay, interest and screening mechanism. Also variables like age, experience, loan inadequacy and loan type are significant in Bayesian analysis in addition to significant predictors in classical logistic regression.

The first factor which affects repayment efficiency is loan size. The availability of sufficient loan size is one important factor. Thus, it is good to compare loan size with the business proposal of the client before loan disbursement and should revise the rule and regulation of the institution based on the current economic condition of the country. The study by: Ojiako and Ogbukwa. (2012), implies that as amount of loan increases, the opportunity to run larger projects increases making them more competitive and profitable. Similarly, the study by: Mokhtar, Nartea and Gan (2010) indicated that; the determinants of loan repayment problems among the Malaysian borrowers showed that loan amount were among the factors that influenced borrowers in repaying their loans. Also, Roslan (2007) & Mullineaux (2009) and et al., Reported similar results.

Monthly income also has positive significant contribution to the repayment efficiency, as income increases then the repayment efficiency also increases more likely than those whose income is not increased. Lehnert, (2004) and Nannyonga (2000) reported that, faster income growth is associated with efficient repayment and low income is associated with inefficient

repayment performance. Ojiako and Ogbukwa, (2012) reported that income has significant contribution for the repayment efficiency.

The educational status of the borrowers is significant in both approaches, which is major factor affecting repayment efficiency of borrowers as many literatures outlined from the economics and business areas. For example, the study by Micha'el(2006) indicates, better repayment performance is strongly and directly associated with educational level of the borrowers. This statements from the assumption that, those who have attended more of formal education than who have not, shall plan and evaluate their business well before taking the credit. In many empirical studies, it was found that more educated beneficiaries tend to use the loan funds for the intended purpose than less educated or non-educated borrowers (Godquin, 2004).

Family size, which is defined as the total number of individuals in the family and elsewhere that depend on the borrower is another factor affecting repayment efficiency. Micha'el (2006), Ojiako and Ogbukwa (2012) reported that household size had a negative influence on the repayment capacity of borrowers i.e. as the number of dependent increases, the borrower will need more money to fulfill their requirements in addition to the obligation of loan repayment. As a result he/she may divert the loan to meet their needs, increases expenditure cost and reduces repayment efficiency.

Suitability of time or duration given to repay loan has significant contribution on the repayment efficiency. If enough time is given for borrowers to repay loan, they can have better repayment efficient than the current two year. Mullineaux (2009) reported that repayment efficiencies are nonlinearly related to the length of time to emergence. Similar study by Jemal (2003) and Donald (2007) reported that “if borrowers find the repayment

period suitable, they can utilize the loan effectively for the intended purpose than those who said the period of repayment is unsuitable”.

Considering sex and age, female and older borrowers were worse loan payers than male and younger borrowers. This can be due to high work load, cultural determination, problem of lack of experience and exposure to business in comparison to male borrowers and as borrower becomes elder, they might be unable to compete with young individuals which is similar with the study by Berhanu (2005) and Godquin (2004), However this does not agree with the econometric result of Jemal (2003).

Additional income, as the presence of other income separated from major income increases, the rate of credit default declines. This would suggest that as clients expand their capital base through increased access to financial services and diversify their sources of income by starting other businesses, then their repayment efficiency can be improved. A woman running a clothing shop for example decided to use her next loan to start trading in cereals just outside her shop. This finding is consistent with a study undertaken among borrowers in Caja Los Andes, Bolivia and Ghana, which indicates that borrowers with many income sources are less likely to default than those having only one income source (Pollio and Obuobie, 2010).

In the case of business experience, as the number of years a borrower has been in business increases, the probability of default declines. This result was supported by Pollio and Obuobie (2010) which stated that, as the number of years a borrower has been in business increases, the probability of default declines by 28 percent. This confirms that as borrowers gain commercial experience, the resulting improved productivity leads to a significant reduction in likelihood of default compared to less experienced counterparts.

When coming to the interest rate the institution receives has a significant contribution for repayment efficiency. Keynesian economists recommended that interest rates should be kept low in order to speed the growth of investment and economy at large (Roe 1982). The virtues of low interest rates are: it will increase borrowing, reduce inflation, increase job opportunities and stimulate national economy. Stiglitz and Weiss (1981) believe that high interest rates are responsible for higher defaults and declining bank profit. These indicate that high interest rates are positively correlated to loan defaults in developing countries.

Variables like: Motivation, screening mechanism, number of reputation, inadequacy of loan and loan type are also significant predictors. Similar study by Jemal (2003) indicates that, repeatedly borrowed customers acquired more experience on the institutional rules, regulations and hence could efficiently utilize the loan for the intended purpose and repay without any difficulty. Also, Pollio and Obuobie (2010) identified decrease with the number of dependents, presence of transparent screening mechanism, frequency of monitoring clients, years in business, the number of guarantors and motivating borrowers are factors associated with repayment efficiency of borrowers. The result from the models, Bayesian analysis predicted the outcome variable well than the result from the classical logistic regression. i.e. from the same variables used in the analysis, 12 variables are significant in classical one and 16 variables are significantly predict outcome variables in Bayesian approach. These can be due to incorporation of prior information in addition to data on hand and availability of sufficient sample size from the simulation than classical logistic regression even if the prior information used in Bayesian analysis is uninformative.

## **10. Conclusion and Recommendation**

### **9.1. Conclusions**

The descriptive analysis of loan efficiency shows that out of 340 borrowers considered, 38.5% were not efficient on repayment and the remaining 61.5 % of them were efficient on repayment at the time of the study period.

The PCFA using principal component method with varimax rotation: Benefit and obstacle related factors like peer effect and experience which account 25% of total variation explained the impact of the efficient utilization of the loan to the borrowers, also good saving habit, high capital accumulation, satisfaction level on the service, improvement on consumption by decreasing expenditure cost has significant impact on the efficient utilization of loan and business success. Thus, by working on those factors, it is possible to improve efficient utilization of loan to see positive impact on the livelihood of borrowers.

Results of classical binary logistic and Bayesian logistic analysis, supporting female borrowers, having proportion family size with income, educating societies, increasing monthly income and loan Size, diversifying source of income, balanced tax system by the government, increasing time given to repay loan, motivation of repayment by different ways, minimizing interest and creating good screening mechanism when borrowers apply to loan have significant impact on loan repayment efficiency of borrowers in the Hawassa city. From these predictors: family size, tax and interest have a negative relationship with outcome variable, whereas monthly income and the rest have positive significant effect on the repayment efficiency of the borrowers. In addition to above predictors: age, Loan type, Purpose of loan, Inadequacy of loan and experience has significant effect on loan repayment efficiency using Bayesian analysis.

## 9.2. Recommendations

This study has found that improving the loan efficiency is a prerequisite to making the business profitable.

- ✓ Strengthen its management information systems to produce up-to-date loan repayment statements for borrowers and to enable early detection of potential default and slower payment problems.
- ✓ Increasing loan size to run business in more competitive manner must have to be given special attention by minimizing interest.
- ✓ MFIs should create such incentives, support and increase the time given to repay loan that would motivate borrowers to repay their loans without any difficulty.
- ✓ MFIs should devise such policies that credit should reach to the low income group.
- ✓ Although continuous follow up and supervision, Benefit and obstacle related factors, capital accumulation, government

## 11. References

- Abebe M. (2011). Determinants of Credit Repayment and Fertilizer Use by Cooperative Members in Ada District: East Shoa Zone, Oromia Region; M.Sc. Thesis, Haramaya University.
- Abreham Gebeyehu (2002). Loan repayment and its Determinants in Small-Scale Enterprises Financing in Ethiopia: A Case of Private Borrowers Around Zeway Area, M. Sc. Thesis, Addis Abeba University.
- Afrin, S. (2008). Multivariate Model of Micro Credit and Rural Women Entrepreneurship Development in Bangladesh: *Bangladesh International Journal of Business and Management*.
- Besag, J. (2001). Markov Chain Monte Carlo for Statistical Inference, University of Washington: USA Center for Statistics and the Social Sciences.
- Grimm, L.G. and Yarnold, P.R. (19195). Reading and Understanding Multivariate Statistics: Washington D.C. Am. Psychol.
- incentive and support, satisfaction on the service, minimize expenditure cost and saving habit is important impact direction of efficient utilization of loan. Thus, the institution should work more in this regard by collaborating with different associations and government.
- ✓ The result also be implemented using classical and Bayesian logit and prohibit models and Bayesian model averaging should be used as they explicitly accounts for model uncertainty and estimates models with every possible combination of regressors and solve problems in low repayment performance of beneficiaries for future.
  - ✓ A longitudinal study is also recommended for future research. The longitudinal study can monitor changes in the borrower's business, household and individual after receiving the microcredit loan.
- Huang, X. (2010). Bayesian Logistic Regression Model for Siting Biomass: Using Facilities.
- Meyer, J.B. (2002). Madness to the Method: Empirically Assessing Small Business Lending Under the Community Reinvestment. *Urban Law Review*, 34 (1), 1-38.
- Micha'el Addisu (2006). Micro-Finance Repayment Problems in the Informal Sector in Addis Ababa: *Ethiopian Journal of Business & Development, Volume 2*.
- Oke, J.T.O. Adeyemo, R. and Agbonlahor, M.U., (2007). An Empirical Analysis of Microcredit Repayment in Southwestern Nigeria, *Humanity & Social Sciences Journal 2* (1): 63-74, ISSN 1818-4960.
- Okorie, A. (1986). Major Determinants of Agricultural Smallholder Loan Repayment in a Developing Economy: Empirical Evidences from Ondo State, Nigeria.

- Okorie, A. (1986). Major Determinants of Agricultural Smallholder Loan Repayment in a Developing Economy: Empirical Evidence from Ondo State Nigeria.
- Okorji, E.C. and Mejeha, W. (1993). The Formal Agricultural Loans in Nigeria: The Demand for Loans and Delinquency Problems of Smallholders Farmers in the Owerri Agricultural Zone of Imo State. *Int. J. Trop. Agric.*, 1:1-13.
- Olagunju, F.I. and Adeyemo, R. (2007). Determinants of Repayment Decision Among Small Holder Farmers in Southwestern Nigeria, *Medwell Pakistan Journal of Social Sciences*.
- Oluwansunmi, A.A. and Aloa, A. (1999). The Role of Credit in the Transformation of Traditional Agriculture: The Nigerian Experience. *Nig. J. Econ. Soc. Studies*.
- Padma, M. and Getachew, A. (2005). Women Economic Empowerment and Microfinance: A Review on Exercises of Awassa Women Clients.
- Pitt, M.M. and Khandker, S.R. (1998). The Impact of Group Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?, *Journal of Political Economy*, Vol. 06, No. 5.
- Podgurky, D., (2000). Determinants of Microcredit Loans Repayment Problem Among Microfinance Borrowers in Malaysia.
- Pollil, G. and Obuobie, J. (2007). Microfinance Default Rates in Ghana: Evidence from Individual-Liability Credit Contracts.
- Roslan, A.H. and Karim, A.Z.A. (2009). Determinants of Microcredit Repayment in Malaysia: A Case of Agrobank. *Humanity Soc. Sci. Journal*; Volume 4: P.45-52.
- Salie Ayalew (2007). Empirical Impact Assessment of Business Development Service on Micro and Small Enterprises in Towns of Amhara National Regional State: M.Sc. Thesis in Addis Ababa University.
- Sean, M.O. and David, B.D. (2007). Bayesian Multivariate Logistic Regression: Biostatistics Branch, National Institute of Environmental Health Sciences.
- Steiner, M. and Tym, C. (2005). Multivariate Analysis of Student Loan Defaulters at the University of South Florida: TG Research and Analytical Services.
- Steiner, M., and Teszler N. (2003). The Characteristics Associated with Student Loan Default at Texas: A and M University.
- Stiglitz, J.E. and Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information, *The American Economic Review*, Volume 71, Issue 3, pp 393-410.
- Sydney, S. (2007). A Bayesian Approach to Variable Selection in Logistic Regression with Application to Predicting Earnings Direction, Ron Bird and Anthony Hall School of Finance and Economics University of Technology.
- Tedeschi, G. A. (2006). "Here today, gone tomorrow: Can dynamic incentives make Microfinance More Flexible?", *Journal of Development Economics*.
- Tesfaye, G.B. (2009). Econometric Analyses of Microfinance Credit Group Formation, Contractual Risks and Welfare in Northern Ethiopia: PhD Thesis, Wageningen University, Netherlands.
- Volkwein, N.H. (1995). The Efficiency of Microfinance in Vietnam: Evidence from Ngo schemes in the North and the Central Regions.
- Woo, B.H. (2009). Women and Repayment in Microfinance: Institute of Research for Development in France, University of Agder, Norway Working Paper.