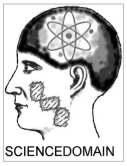
British Journal of Mathematics & Computer Science
19(2): 1-15, 2016; Article no.BJMCS.23206
ISSN: 2231-0851

SCIENCEDOMAIN *international*
www.sciencedomain.org



Bayesian and Classical Approaches Applied on Saving Habits of Employees Data: A Comparative Study

Genanew Timerga[1] and Nigatu Degu[2*]

[1]Department of Statistics, Debre Birhan University, Debre Birhan, Ethiopia.
[2]School of Mathematical and Statistical Sciences, Hawassa University, Hawassa, Ethiopia.

Authors' contributions

This work was carried out in collaboration between the two authors. Both authors involved in managing the data, carrying out the statistical analysis, designing the study, writing the protocol, finalizing the manuscript and literature searches. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMCS/2016/23206
Editor(s):
(1) H. M. Srivastava, Department of Mathematics and Statistics, University of Victoria, Canada.
Reviewers:
(1) P. E. Oguntunde, Covenant University, Nigeria.
(2) Diana Bílkova, University of Economics, Prague, Czech Republic.
(3) Rashmi Awad, Devi Ahilya University, Indore, India.
Complete Peer review History: http://www.sciencedomain.org/review-history/16638

Original Research Article

Received: 19th November 2015
Accepted: 15th December 2015
Published: 22nd October 2016

Abstract

480 employees questionnaire were collected on saving habit at Debre Birhan town during February, 2010 to October, 2011. Descriptive statistics, Binary logistic regression and Bayesian statistical methods were used. Average private worker, low government worker were associated with a significantly lower likelihood of saving regularly versus not saving. Binary logistic regression indicates that age, education, dependent family members, transport, job satisfaction, expenditures and inflation significantly affect saving habits of employees. (Coeff-0.569, OR 0.566, P=0.000, CI 0.468, 0.685) the odds of saving decreases by 43.4% for one unit increase in dependent family members. The regression coefficient for the consumption growth rate on the one-period lagged consumption growth rate is expected to be positive. Capacity of employees' utilized, formal method of saving institutes is higher than informal saving institutes. Our measures that are expected to capture various precautionary saving habits, that is, number of earners in a family and job security of the head of household, are not perfect in capturing uncertainty about future income.

Keywords: Logistic regression; Bayesian analysis; saving habit; MCMC methods.

Corresponding author: E-mail: ndegu9@gmail.com;

1 Introduction

The goal of promoting financial saving habit is to make people more aware of financial opportunities, choices, and possible consequences. There is a growing recognition of the importance of financial education in relation with saving [1,2]. Moreover, financial education is one way of increasing savings and asset accumulation. According to [3] the extent to which an individual understands the process and benefits of asset accumulation is likely to affect their willingness to save.

In addition, personal saving has two primary functions. The first one is, it provides the economic security of a safety net, which means, individuals are prepared to face unexpected and irregular financial circumstances through transferring resources from the present to the future savings. The other is, saving leads to accumulation of wealth that enables individuals to improve their living standard and to respond a new opportunities [4]. School, food, and medical expenses, family support, employment creation, income generation are important causes for saving [5].

Most of the developing countries have low rate of saving habits; therefore, improving saving habit is a primary goal for people living in this part of the world [6]. Improving saving habit of individuals is given attention to look at a variety of savings services used by people/customers in the community.

Moreover, people in general and the poor in particular might not have sufficient knowledge with regard to saving [7]. However, in Ethiopia, very limited research works are available in this regards. Therefore, assessing employees' saving habits has been crucial; to enhance the individual living standard, domestic investment, minimization of economic inflation and increase overall GDP of the country.

The aim of this study was designed to achieve the following objectives; to assess the current status of saving practices and factors affecting saving habits of the employees at Deber Berhan town.

2 Materials and Methods

The study was conducted at Debre Birhan town, the capital of North Shew zone of Amhara Region from February to October, 2011. The study was targeted at all employees at Debre Berhan town. A total of 5113 employees within the age interval of 18 to 75 years were considered.

The cross-sectional data were collected using a structured questionnaire. Stratified sampling method were employed, government (Strata I) and private employees (Strata II) were systematically selected based on their probability proportions. The sampling frame was prepared from government and private employees by giving sequential order number.

2.1 Sample size determination

The sample size was determined based on stratified sampling with proportions of 95 percent confidence level. Moreover, three percent of the sample size is added to compensate for non-response rate. The sample size formula is given by [8].

$$n = \frac{\sum_{h=1}^{2}\left({W_h}^2 N_h PQ \Big/ W_h(N_h - 1)\right)}{V + \frac{\sum_{h=1}^{2}\left(\frac{W_h N_h PQ}{N_h - 1}\right)}{N}} \quad (1)$$

Where Z be the upper $\alpha/_2$ point of standard normal distribution, where $\alpha = 0.05$ significance level, which is $Z_{\alpha/_2} = Z_{0.025} = 1.96$. Suppose the relative error d is usually set by the investigator desired which is used

from a similar study undertaken by the "Saving habits, needs and priorities in urban Uganda" in 2005, with a sample size of 852 is taken as a reference for the purpose of fixing the standard deviation. One of the explanatory variables used in that study is customer saving with a standard deviation of 0.45. This helps us to determine the sample size to represent the population by calculating the acceptable absolute error, d, where $d = Z\alpha_{/2} S/\sqrt{n}$. If we adopt a significance level α = 0.05, then the calculated margin of error from the above information is 0.03. In our study, it is indicated that the proportion of saving habits of employees p was determined small proportion from the results of previous work of similar population of saving habits of employees in Chinese and Americans 14.02% (p = 0.1402) of [9]. Thus, the sample size calculation is

$$n = \frac{\sum_{h=1}^{2}\left({W_h}^2 N_h PQ \Big/ W_h(N_h - 1)\right)}{V + \frac{\sum_{h=1}^{2}\left(\frac{W_h N_h PQ}{N_h - 1}\right)}{N=(N_1 + N_2)}} = \frac{0.12059}{0.00023 + 0.000024} = \frac{0.12059}{0.000254} \quad = 474.76 \approx 475$$

Finally, 3 percent of the sample size, which is 15, was added to the determined sample size 475 to compensate for none response rate. Thus, the required sample size for this study is 490 employees from 5113 who live at Debre Birhan town. Next, we carried out sample size allocation to each stratum with proportional allocation.

$$\mathrm{k_h} = \left[\frac{\mathrm{nN_h}}{\mathrm{N}}\right] = \mathrm{nW_h} \ , \mathrm{h} = 1{,}2 \tag{2}$$

Where, h = types of employee = 2

$$\mathrm{n} = \sum_{\mathrm{i}=1}^{2} \mathrm{k_h} = 474.76 \approx 475 + 15 = 490$$

$\mathrm{W_h} = {\mathrm{N_h}}/{\mathrm{N}}$ = Probability of stratum weight selection of h^{th} employee, N_h = Total number of employees within the h^{th} strata, N = Total number of employee in DB town

Table 1. Number of employee taken from the selected PSU at Debre Birhan

Employee	N_h	W_h	Sample
Government (N_1)	3503	0.685	336
Private (N_2)	1610	0.315	154
Total	5113		490

2.2 The study variables

2.2.1 Dependent variable

The response variable in this study is the status of saving habits at Debre Birhan town. The habits of employees are identified as either save out of income or no save out of income. The response variable is a dichotomous category and my interest of the study is no saving out of income. Thus, coded as the value 0 for 'save out of income' and 1 for 'no save out of income'.

2.2.2 Independent variables

The independent variables in this study are classified as occupational variables, economic variables and personal relationship and contextual variables.

2.2.3 Logistic regression model

Logistic regression analysis extends the techniques of multiple regression analysis to research situations in which the outcome variable is categorical with two categories. The response variable of the study is binary, (save and no save). Thus, in logistic regression will be appropriate.

2.2.4 Model description

Since the response variable in logistic regression is usually dichotomous, we define such a response variable as Y, and denote the even y=1, when the subject has the characteristic of interest and y=0, when the subject does not have that characteristic of interest. So an alternative form of the logistic regression equation is the logit transformation of P_i given as

$$\text{logit}\left[P_i\right] = \log\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} \tag{3}$$

2.3 Parameter estimation for logistic regression

The maximum likelihood and non-iterative weighted least squares are the two most computing estimation methods used in fitting logistic regression model [10].

Consider the logistic model $P(X_i) = \frac{e^{X\beta}}{1+e^{X\beta}}$, since observed values of Y say, Y_i's (i=1, 2 …, n) are independently distributed as binomial and, the maximum likelihood function of Y is given by:

$$L(\beta|y) = \prod_{i=1}^{n} p(y_i|X_{i1}, X_{i2}, \ldots, X_{ik}) = \prod_{i=1}^{n}\left[\frac{e^{X\beta}}{1+e^{X\beta}}\right]^{y_i}\left[\frac{1}{1+e^{-X\beta}}\right]^{(1-y_i)} \tag{4}$$

Lastly, we check Model Selection, goodness of Fit of the Model, likelihood-Ratio Test, the Hosmer and Lemeshow Test Statistic, the Wald Statistic, R^2 Statistic

2.4 Bayesian analysis

2.4.1 Bayesian logistic regression

The foundation of Bayesian statistics is the Bayes' theorem which states that if A and B, are events and $\boldsymbol{P}(\boldsymbol{B})$, the probability of event B, is greater than zero, then:

$$P(A/B) = \frac{P(A)P(B/A)}{P(A)}$$

2.5 The likelihood and prior function

2.5.1 Likelihood function

The likelihood function of the Bayesian formulation for the joint distribution of n independent Bernoulli trials was still the product of each Bernoulli distribution, the sum of independent and identically distributed Bernoulli trials in which the sum has a Binomial distribution. Specifically, let $y_1, y_2, \ldots, y_n$ be independent Bernoulli trials with success probabilities $P_1, P_2, P_3, \ldots, P_n$, that is $y_i = 1$ with probability P_i or $y_i=0$ with probability $1- P_i$, for i= 1,2,…,n. The trials are independent, the joint distribution of $y_1, y_2, \ldots y_n$ is the

product of n Bernoulli probabilities. As usual, the likelihood, function used by Bayesians matches that from frequents inference. $L(y \mid \beta) = \prod_{i=1}^{n} \left[{P_i}^{y_i} (1 - P_i)^{(1-y_i)} \right]$

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.

$$P_i = \frac{e^{\beta_0+\beta_1 x_{i1}+.....+\beta_k x_{ik}}}{1+e^{\beta_o+\beta_1 x_{i1}+....+\beta_k x_{ik}}}$$

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

$$L(y \mid \beta) = \prod_{i=1}^{n} \left(\frac{e^{\beta_o+\beta_1 x_{i1}+....\beta_k x_{ik}}}{1+e^{\beta_o+\beta_1 x_{i1}+.....\beta_k x_{ik}}} \right)^{y_i} \left(1 - \frac{e^{\beta_o+\beta_1 x_{i1}+....+\beta_k x_{ik}}}{1+e^{\beta_o+\beta_1 x_{i1}+.....\beta_k x_{ik}}} \right)^{(1-y_i)}$$

2.5.2 Prior function

The choice of an informative prior distribution typically involves a certain amount of subjectivity; historically, this has been a reason for disagreement between Bayesian and classical statisticians. The assumed prior normal distribution for parameter β_j is given by

$$f(\beta_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{1}{2}\left(\frac{\beta_j - \mu_j}{\sigma_j} \right)^2 \right) \tag{5}$$

2.5.3 The posterior distribution

Based on 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

$$f(\beta|y) = \prod_{i}^{p} \frac{{p_i}^{y_i}(1-p_i)^{1-y_i}p(\beta)}{p(x_1, x_2, \dots, x_p)} \propto \prod_{i}^{p} {p_i}^{y_i}(1-p_i)^{1-y_i}p(\beta)$$

$$f(\beta|y)$$

$$= \prod_{i=1}^{p} \left[\left(\frac{e^{\beta_0+\beta_1 X_{i1}+\cdots+\beta_p X_{ip}}}{1+e^{\beta_0+\beta_1 X_{i1}+\cdots+\beta_p X_{ip}}} \right)^{y_i} \left(1 - \frac{e^{\beta_0+\beta_1 X_{i1}+\cdots+\beta_p X_{ip}}}{1+e^{\beta_0+\beta_1 X_{i1}+\cdots+\beta_p X_{ip}}} \right)^{1-y_i} \right]$$

$$\times \prod_{i=0}^{p} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\left(\frac{\beta_i - \mu_i}{\sigma_i} \right)^2 \right) \tag{6}$$

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

2.6 Markov Chain Monte Carlo (MCMC) methods

The use of Markov chain Monte Carlo (MCMC) methods to evaluate integral quantities has exploded over the last fifteen years. The primary distinction made here is between standard Monte Carlo simulation and the Markov chain type of Monte Carlo methods.

2.7 The Gibbs sampler algorithm

The Gibbs sampler [11] is the most widely used MCMC technique. If the limiting distribution of interest is $\pi(\beta)$ where β is an k length vector of coefficients to be estimated, then the objective is to produce a Markov chain that cycles through these conditional statements moving toward and then around this distribution. The set of full conditional distributions for β are denoted β and defined by $\pi(\beta) = \pi(\beta \mid \beta_i)$ for i = 1, 2… k, where the notation β_i indicates a specific parametric form from β without the β_i coefficient described as:

1. Start with an initial value: $\beta^{[0]} = \beta^{[0]}{}_0, \beta^{[0]}{}_1, \beta^{[0]}{}_2, \ldots, \beta^{[0]}{}_k$
2. Sample for each i = 0, 1, 2... , n-1:

   Generate $\beta_0^{(i+1)}$ from $f\left(\beta_0 \mid \beta_1^{(i)}, \beta_2^{(i)}, \beta_3^{(i)}, \ldots, \beta_k^{(i)}\right)$

   Generate $\beta_1^{(i+1)}$ from $f\left(\beta_1 \mid \beta_0^{(i+1)}, \beta_1^{(i)}, \beta_2^{(i)}, \ldots, \beta_k^{(i)}\right)$

   Generate $\beta_2^{(i+1)}$ from $f\left(\beta_2 \mid \beta_0^{(i+1)}, \beta_1^{(i+1)}, \beta_2^{(i)}, \ldots, \beta^{(i)}{}_k\right)$

   ⋮

   Generate $\beta_k^{(i+1)}$ from $f\left(\beta_k \mid \beta_0^{(i+1)}, \beta_1^{(i+1)}, \beta_2^{(i+1)}, \ldots, \beta_{k-1}^{(i+1)}\right)$ ..... .(3.16)
3. Return $\beta^{(1)}, \ldots\ldots\ldots\ldots, \beta^{(n)}$

Once convergence is reached, all simulation values are from the target posterior distribution and a sufficient number should then be drawn so that all areas of the posterior are explored.

3 Results and Discussion

3.1 Demographic characteristics

As indicated in Table 2, about 47.3% employees had good saving habits; however; the remaining 52.7% did not practice saving at the time of data collection. Out of the total sampled population, 31.5% of the employees were in private organization whereas 68.5% in government organization. With regard to the sex composition, 43.7% were female and 56.3% were male employees. The age distribution indicates that 26% of the respondents were in the age category below 25 years, 37.4% in 25-35 years, 20.8% in 36-45 and 15.8% were above 45 years old. In case of marital status, 44%, 52.1%, 3.8% and 0.1% of the respondents were single, married, separated and widowed, respectively. Majority of employees (83.1%) were Orthodox follower followed by Muslim (4.2%), Protestant (9.4%), and other religion (3.3%). Almost all of the respondents (82.9%) their monthly salary income was above 1000 Ethiopian birr, where as 17.1% were below 1000birr. With regard to educational status, 88.5% of the employees were above diploma holders. From cross-tabulation, it can be shown that 61.9% among male and 41% among female employees practices saving. This indicates that proportion of male employees who practiced saving is higher than that of female employees.

Table 2. Results of demographic characteristics (Debre Birhan Town, 2011)

Explanatory variables	Current status of saving habits at DB in 2011					Total	
	Category	No		Yes		Count	%
		Count	%	Count	%		
Sex of the respondents	Female	124	59.0	86	41	210	43.7
	Male	103	38.1	167	61.9	270	56.3
Age of the respondents	Below 25	74	59.2	51	40.8	125	26.0
	25—35	87	48.6	92	51.4	179	37.4
	36—45	41	41	59	59	100	20.8
	Above 45	25	32.9	51	67.1	76	15.8
Education level	Certificate & below	37	67.3	18	32.7	55	11.5
	Diploma	105	58.7	74	41.3	179	37.3
	1st degree	78	38.8	123	61.2	201	41.8
	Masters and above	7	15.6	38	84.4	45	9.4
Religion	Orthodox	189	47.4	210	52.6	399	83.1
	Muslim	8	40	12	60	20	4.2
	Protestant	23	51.1	22	8.7	48.9	9.4
	Other	7	43.8	9	56.2	16	3.3
Ethnicity	Oromo	17	36.2	30	63.8	47	9.8
	Amhara	185	50.4	182	49.6	367	76.4
	Tigre	11	36.7	19	63.3	30	6.3
	Gurage	6	46.2	7	53.8	13	2.7
	Other	8	34.8	15	65.2	23	4.8
Monthly salary	Below 1000	53	64.6	29	35.4	82	17.1
	1000—2000	103	52	95	48	198	41.3
	2001—3000	48	42.5	65	57.5	113	23.5
	Above 3000	23	26.4	64	73.6	87	18.1
Marital status	Single	106	50.2	105	49.8	211	44.0
	Married	109	43.6	141	56.4	250	52.1
	Separated	12	66.7	6	33.3	18	3.8
	Widowed	0	0.0	1	100	1	0.1
Employee	Government	167	50.8	162	49.2	329	68.5
	Private	60	39.7	91	60.3	151	31.5

Table 3 showed that the major reasons employees had not practice saving are low monthly salary income (84%), long distance of financial institution (11%), low interest rate earned (53%) and limited awareness of saving (77%). From those individuals who have saving habits, 86.2% of the respondents have used formal method of saving institutes whereas 75.1% of the respondents have used informal method of saving institutes. This shows that formal method of saving institutes is slightly higher than informal saving methods of saving institutes (Table 3).

Table 3. Reasons for not saving and methods of saving of employees

Variables	Category	Count	%
Saving methods	Formal institutes	218	86.2
	Semi formal institutes	102	40.3
	Informal institutes	190	75.1
Reasons for not saving	Low income in the job	191	84.1
	Long distance financial institution	25	11
	Low interest rate earned	106	46.7
	Limited awareness of saving benefits	52	22.9

3.2 Multiple logistic regression analysis

Forward multiple logistic regression analysis is carried out, and only 10 the most important variables were found to have significant effect on saving practices at 5% level of significance (Table 6).

3.3 Assessing model fit

After the logistic model, selected predictor variables in the forward likelihood ratio selection procedure, the first step is to assess the overall fit of the model to the data. The Hosmer and Lemeshow goodness of fit test divides cases into deciles based on predicted probabilities (Table 4). The SPSS output in Table 4 shows the non-significance of the chi-square value. Hence, we do not reject the null hypothesis; there is no difference between the observed and expected frequencies which indicates that the model adequately fits the data.

The omnibus and chi-square tests are measures of how well the model performs and the difference between the initial model and the regression model in terms of number of correctly classified subjects or it is the change in the -2log-likelihood from the previous step respectively, final step is considered be appropriate (Table 4).

Table 4. Results of omnibus tests of model coefficients

	Final step	Chi – square	df	Sig.
		6.160	8	0.629
	Step	7.694	1	0.006
Final step of	Block	207.548	19	0.000
omnibus tests	Model	207.548	19	0.000

The model summary with -2log-likelihood statistic shows the overall fit of the model. Cox and Snell R square is 0.351, that is 35.1% of the variation in the dependent variable is explained by the predictor variables and the nagelkerke R-square shows that approximately 47% of in the dependent variable is explained by the predictor variables (Table 5).

Table 5. Likelihood and pseudo R square

Final step	-2log likelihood	Cox and Snell R square	Nagelkerke R square
	456.464	0.351	0.469

Since, most of the covariates are categorical to compute odds ratio we need to have a reference category. Multiple logistic regression coefficients can be estimated using the maximum likelihood estimation method implemented in the SPSS package (Table 6).

The estimated coefficients and standard errors of the estimates that are used in computing the Wald statistic and the odds ratio (Exp(β)) are presented in Table 6. The significance of the Wald statistic indicates the importance predictor variable in the model.

3.4 Discussion

The study shows that number of dependent family member has a significant effect on saving habit of the employees. The odds ratio indicates that one number increase of dependent family member, the probability of not to save of the same employee increases 0.566 times that of an employee who has no dependent family member (coeff -0.569, OR 0.566, P=0.000, CI 0.468, 0.685). Employees with large number of dependent family members are more likely not to save money as compared with small number of dependent family

members. This indicate that after controlling other variables in the model, the odds of saving is decreased by 43.4% for each unit increases in number of dependent family members.

Table 6. Final results of multiple logistic regression models

Variables	Category	β	S.E.	Wald	df	Sig.	Exp(β)	95% CI for EXP (β)	
								Lower	Upper
Age	Below 25 (Ref.)			33.076	3	.000			
	25—35	.603	.294	4.202	1	.040	1.828	1.027	3.253
	36—45	1.409	.383	13.533	1	.000	4.094	1.932	8.674
	Above 45	2.484	.444	31.247	1	.000	11.991	5.019	28.649
Education	Certificate & below (Ref.)								
	Diploma	.347	.396	.767	1	.381	1.414	.651	3.071
	1st degree	.763	.403	3.581	1	.058	2.144	.973	4.724
	Masters & above	2.265	.652	12.065	1	.001	9.628	2.683	34.552
Housing	Owner (Ref.)			11.799	2	.003			
	Rent private	.908	.271	11.241	1	.001	2.478	1.458	4.213
	Rent Government	.820	.485	2.858	1	.091	2.271	.878	5.875
Member of saving no (Ref.)									
	Association Yes	1.273	.268	22.557	1	.000	3.571	2.112	6.039
Transport service no (Ref.)									
	Yes	.793	.291	7.444	1	.006	2.211	1.250	3.909
Job satisfaction in the sector	Unsat. (Ref.) Satisfied								
		.913	.286	10.167	1	.001	2.491	1.421	4.365
Inflation affect saving	No (Ref.) Yes	1.639	.543	9.115	1	.003	5.152	1.777	14.935
Cost of expenditures									
	Below 1000 (Ref.)								
	1000—1500	.100	.306	.106	1	.745	1.105	.606	2.014
	1501—2500	.497	.318	2.433	1	.119	1.643	.880	3.067
	Above 2500	-1.712	.640	7.157	1	.007	.181	.051	.633
Cost of recreation									
	Below 100 (Ref.)								
	100—250	.922	.311	8.813	1	.003	2.514	1.368	4.621
	251—350	1.202	.411	8.548	1	.003	3.327	1.486	7.450
	Above 350	1.397	.461	9.197	1	.002	4.042	1.639	9.968
	Constant	-4.265	.779	29.989	1	.000	.014		

Ref. = Reference group
Source: Own research

Housing is another significant factor on the saving habit of employees. The odds ratio of "not to save" for employees who rent private house was 2.478 times higher compared to house owner employees (coeff 0.908, OR 2.478, P=0.001, CI 1.458, 4.213). This result indicates that employees who rent private house have low saving habits than house owners. This result similar with the study conducted in New Zealand by [12] which reported saving habits of house owners is more than those who rent house.

Education status of employees was also found to affect the saving habit of employees. The probability of "not to save" for masters and above is 9.628 times compared to those whose education status was certificate

and below (coeff 2.265, OR 9.628, P=0.001, CI 2.683, 34.552). Similar Studies conducted by [13] and [14] showed that highly educated individuals tend to have higher average saving habits.
Finally, saving association member has also impact on saving habit of employees. Those who were membership of saving association have 3.571 times more chance to save compared to those who are not membership of saving association (coeff 1.273, OR 3.571, P=0.000, CI 2.112, 6.039). [12] and [15] indicate that being a membership of saving association improves saving habits of employees. Other Significant variables are interpreted with similar manner.

3.5 Bayesian logistic analysis

For Bayesian logistic regression analysis, Gibbs sampling method in the WinBUGS was conducted. The Bayesian model used is normal-normal, in which the dependent variable saving habits is assumed to follow a normal distribution with the prior of the coefficients 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 $\boldsymbol{\sigma^2}$ with shape parameters 0.01. Among the variables considered the number of family member, number of dependent family member and distance of working place are continuous variables. Three chains of parameters were simulated for 50000 iterations each. Total of 30000 posterior samples and the first 20000 iterations are discarded as the burn-in stage by checking the time Series of all the parameters.

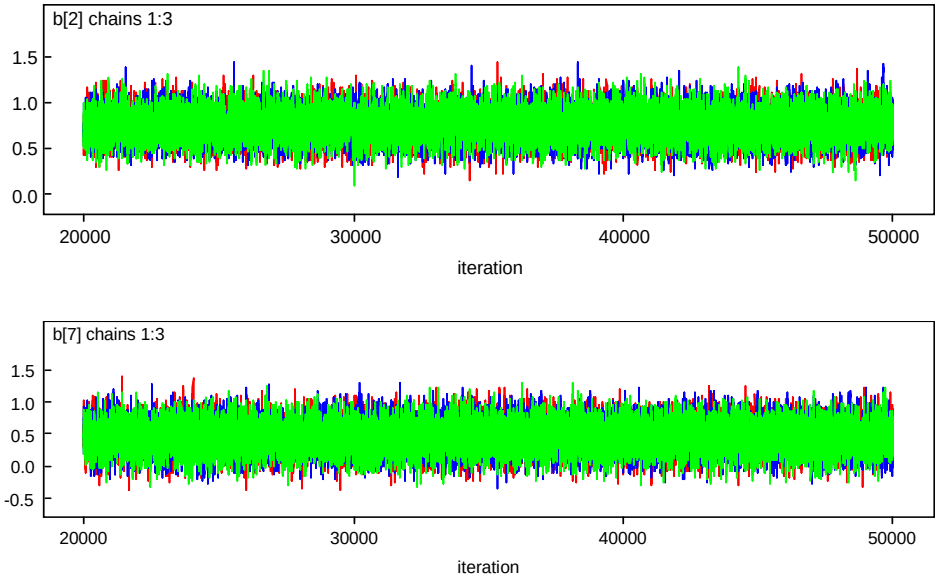
3.6 Assessment of model convergence

Checking the convergences of MCMC output, four methods were used in this study.

3.7 Time series

Time series plots (iteration number on x-axis and parameter value on y-axis) are commonly used to assess convergence. If the plot looks like horizontal band, with no long upward or downward trends, then we have evidence that the chain has converged.

Four independently generated chains demonstrated good “chain mixture” an indication of convergence (Fig. 1). The Time series plots show that the chains with three different colors overlap one over the other.



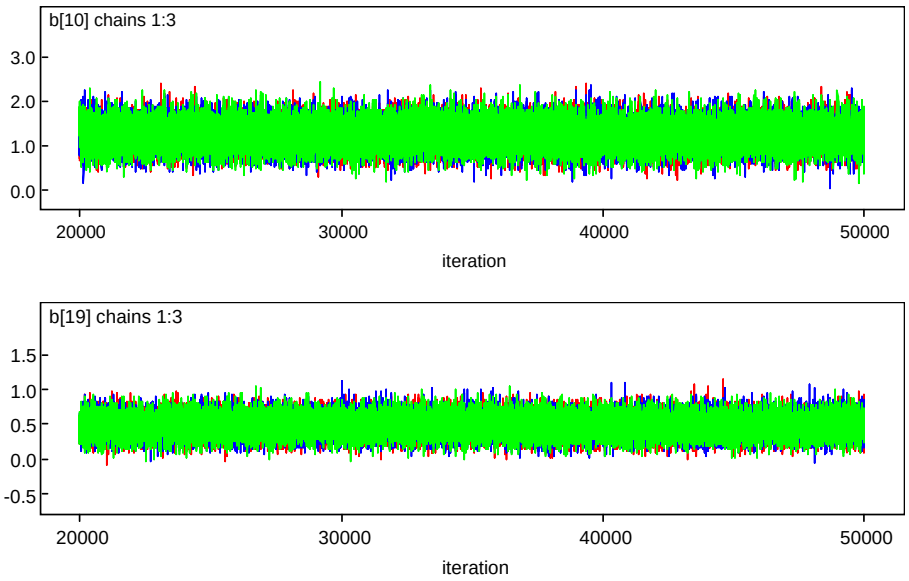


Fig. 1. Convergence of time series plots for the coefficients of age of the respondent, housing, saving association membership, and recreation cost

Source: Own research

3.8 Gelman-Rubin

The Gelman-Rubin statistic is used for assessing convergence. The model is judged to have converged if the ratio within variability is close to 1. Evidence for convergence comes from the red line being close to 1 on the y-axis and the blue and green lines being stable across the width of the plot. Since in our plot the red line is close to one, we can consider this is evidence for convergence.

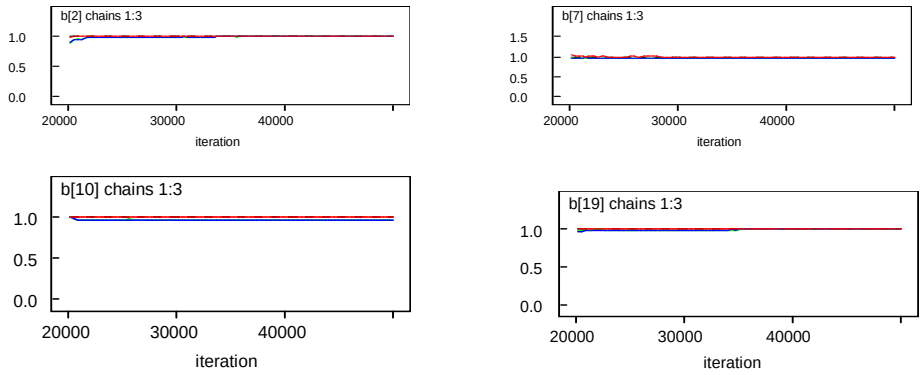


Fig. 2. Convergence using Gelman-Rubin statistic for the coefficients of age of the respondent, housing, saving association membership, and recreation cost

Source: Own research

3.9 Autocorrelation function

High autocorrelations indicate slow mixing within a chain, usually slow convergence to the posterior distribution. The plots show that the four independent chains were mixed or overlapped to each other and hence this is an evidence of convergence.

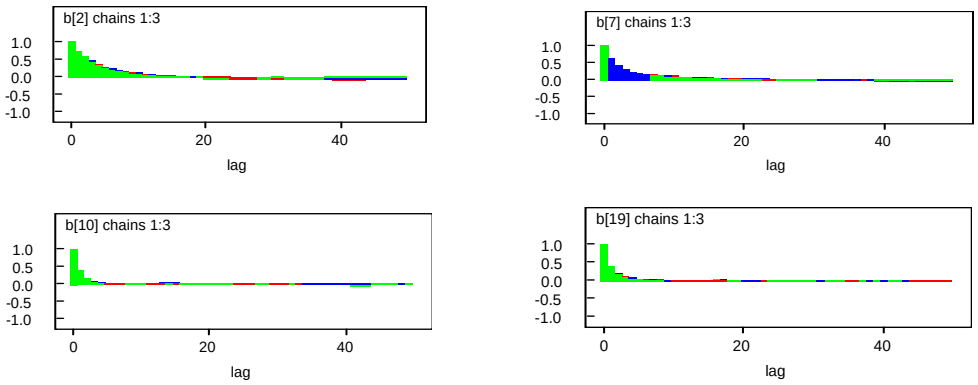


Fig. 3. Convergence of autocorrelation plots for the coefficients of age of the respondent, housing, saving association membership, and recreation cost

3.10 Kernel density

Density used as alternative method for identifying model convergence. The plots for all statistically significant covariates indicated none of the coefficients have bimodal density, and hence the simulated parameter values were converged.

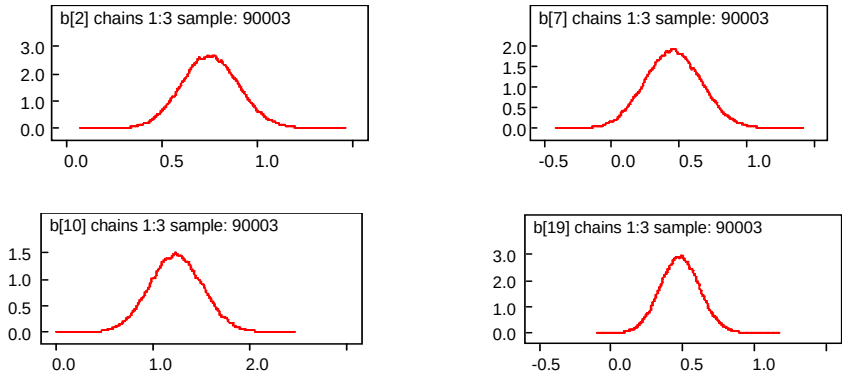


Fig. 4. Convergence for density plot of the Parameter's of age of the respondent, housing, saving association membership, and recreation cost

Table 7 showed that, 95% credible interval for intercept of monthly salary, the posterior standard deviation σ excludes 0. The mean value related to "not to saved" with monthly salary less than 1000, whose posteriori expected to be equal to β_6=0.1479 with a standard deviation of 0.1767. Comparing the posterior mean of the parameter, supporting others with money β_{20}= 0.334 with sd=0.2478 than employees who are not supporting others.

95% credible set of distance from home to work place negative values (β_{13}= -0.1275) with CI = (-0.3158, -0.0593), monthly salary β_6 =0.1479 with CI = (-0.2019, -0.4935) and supporting others with money β_{20}=0.334 with CI = (-0.1488, -0.8245) relations to saving habits of employees respectively.

3.11 Model of the Bayesian logistic regression

Based on binary logistic regression analysis and Bayesian logistic regression analysis the following model was written as follows:

$$\ln\left(\frac{p}{1-p}\right) = -4.683 + 0.7537 \times AG + 0.4721 \times ED + 0.1479 \times MS + 0.455 \times HO$$
$$- 0.6898 \times NDF + 1.243 \times MSA - 0.1275 \times DIS + 0.6651 \times TRS + 0.8164 \times JS$$
$$+ 1.761 \times IN + 0.4793 \times REC + 0.334 \times SUP - 0.1041 \times EXP$$

Table 7. Results of Bayesian model

Model parameters	Node	Posterior mean	Standard deviation	MC error	95% credible set	
Alpha		-4.683	0.8701	0.02222	-6.399	-3.012
Sex of the respondents	b_1	0.3478	0.2654	0.002018	-0.1736	0.8702
Age of the respondents *	b_2	0.7537	0.1502	0.001222	0.4648	1.053
Level of education *	b_3	0.4721	0.1905	0.002214	0.1053	0.852
Employer	b_4	0.4533	0.2813	0.001894	-0.09576	1.008
Number family member	b_5	0.0519	0.1239	0.002209	-0.1858	0.2971
Monthly Salary **	b_6	0.1479	0.1767	0.001919	-0.2019	-0.4935
Housing *	b_7	0.455	0.2104	0.001745	0.0502	0.8711
Parent asset on saving	b_8	-0.204	0.2672	0.002074	-0.7304	0.3183
Number of dependent family *	b_9	-0.6898	0.1571	0.002478	-1.005	-0.3865
Member of saving association *	b_{10}	1.243	0.2739	0.001531	0.7127	1.785
Years live at Debre Birhan town	b_{11}	-0.0102	0.11	7.944E-4	-0.2236	0.2053
Saving money for quality of life	b_{12}	0.4249	0.3402	0.004269	-0.2405	1.088
Distance home to work place **	b_{13}	-0.1275	0.09564	7.163E-4	-0.3158	-0.0593
Transport service in working area *	b_{14}	0.6651	0.3145	0.001675	0.05394	1.284
Getting any allowance service	b_{15}	0.2497	0.2691	0.001219	-0.2778	0.7797
Job Satisfaction in the sector *	b_{16}	0.8164	0.3251	0.003534	0.1773	1.452
Income satisfaction on the job	b_{17}	0.00118	0.268	0.001945	-0.5215	0.5282
Inflation *	b_{18}	1.761	0.5711	0.01229	0.6938	2.935
Recreation cost *	b_{19}	0.4793	0.1375	8.57E-4	0.2144	0.7542
Supporting others with money **	b_{20}	0.334	0.2478	0.001603	-0.1488	-0.8245
Expenditure cost *	b_{21}	-0.1041	0.1564	0.001218	0.4122	0.2001

** Indicates significance in Binary logistic regression and Bayesian analysis*
*** Indicates significance in Bayesian analysis*
Source: Own research

4 Conclusions and Recommendations

4.1 Conclusions

The demographic characteristics of saving habits show that 47.29% of employees have no saving habits and 52.71% have saving habits at the study period. The binary logistic regression and Bayesian logistic analysis showed that age, education, housing, number of dependent family member, saving association member, transport service, job satisfaction, inflation, expenditures and recreation cost were the major factors that affect the saving habits of employees in the town. Family size and lower income are also indicators of low saving habits of employees. The capacity of institutions, employees' utilized formal method of saving institutes is higher than informal saving methods of saving institutes.

Monthly salary, distance of work place and supporting others with money are significant predictor identified using Bayesian analysis. Based on this, number of dependent family, distance of work place and expenditure are correlated negatively with the parameter of saving habits of employees.

4.2 Recommendations

To maximize saving habits of employees in the town, education status should be considered when an employee has a complicated living condition. Appropriate monthly salary and housing should be given special attention.

Putting the above consequences of no saving habits, the following recommendation should be implemented by the concerned bodies:

- Since education status is one of the problems identified in this study, attention should be given education and training of employees.
- Above half percent of employees are private house rents. This reduces the saving habits of the employees. Thus, government should give attention for those people to solve this problem.
- Extended working time has been observed, and as a result employees' satisfaction to their work is low. Hence, add allowance and increased monthly salary for the employees.
- Work more on awareness creation, Support and promote saving associations and advocacy on saving habit.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Greenwald, Grinstein-Weiss, Zhan M, Sherraden M. Saving performance in individual development Accounts: Journal of Marriage and Family. 2001;192-204.

[2] Gill. Voicing the silent fear: South Asian women's experiences of domestic violence. The Howard Journal of Criminal Justice. 2004;43(5):465–483.
DOI: 10.1111/j.1468-2311.2004.00343.x

[3] Johnson E, Sherraden MS. From financial literacy to financial capability among youth. Journal of Sociology and Social Welfare. 2007;34(3):119-145.

[4] Gokhale. Personal saving in developing nations, further evidence. Private Pensions and Public Policies. Brookings Institution Press. Amazon.com; 2000.

[5] Kassa Woldesenbet. Impact analysis of southern ethiopia micro enterprise project on beneficiaries, social & economic. A Paper Presented to the National Micro Enterprise Work Shop (Amharic Version), Addis Ababa; 1998.

[6] Michael. Model selection: Full Bayesian approach. Wiley Online Library; 2001.
DOI: 10.1002/env.482

[7] Karim M. Participation in private retirement savings plans, 1997 to 2008. MPRA Paper 22711, University Library of Munich, Germany; 2010.

[8] Contrast of Bayesian and classical sample size determination. Journal of Modern Applied Statistical Methods. 2014;13:2.

[9] Fei-Fei L. Bayesian variable selection in structured high-dimensional covariate spaces with applications in genomics. Journal of the American Statistical Association. 2010;105(491):1202-1214.

[10] Hosmer D, Lemeshow S. Applied logistic regression. New York: John Wiley and Sons, Inc; 1989.

[11] David M. Variational inference for Dirichlet process mixtures. Bayesian Analysis. 2006;1(1):121-144.

[12] Andrew, Coleman. Inflation and the measurement of saving and housing affordability. Motu Economic and Public Policy Research: Motu Working Paper 08-09; 2008.

[13] Beverly S, Clancy M. Financial education in a children and youth savings account policy demonstration: Issues and options (research background paper CYSAPD 01-5). St. Louis: Washington University, Center for Social Development; 2001.

[14] Solomon I. Random Walks in a random environment. Ann. Probab. 1975;3(1):1-31.

[15] Masson. International evidence on the determinants of private saving. World Bank Econ Rev. 1998;12(3):483-501.
DOI: 10.1093/wber/12.3.483

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/16638