Bayesian Analysis of Retention and Graduation of Female Students of Higher Education Institution: The Case of Hawassa University (HU), Ethiopia

Tsega Kahsay Gebretekle[*], Ayele Taye Goshu

Department of Mathematics, Kotebe Metropolitan University, Addis Ababa, Ethiopia

Email address:
tsegaka@yahoo.com (T. K. Gebretekle), ayele_taye@yahoo.com (A. T. Goshu)
[*]Corresponding author

Abstract: The study was conducted on female students who were 2005, 2006, 2007, and 2008 entries in the fields of Natural Science, Agriculture, and Social Science. From 1931 female students a sample of 605 was taken using stratified random sampling, Primary and secondary data were collected using questionnaire and analyzed using the Bayesian logistic regression analysis. The results showed that the percentage of graduation among 362 females who were enrolled in 2005, 2006, and 2007 was 72.1%. Similarly the retention rate among 243 females of 2008 entry was 75.7%. From the Bayesian logistic regression analyses, significant predicators of both graduation and retention were choice of field, preparatory average result, entrance exam score and first year cumulative GPA. Moreover pregnancy, organizing studying and leisure time, habit of chewing Khat, satisfaction with instructors, parent income, habit of smoking cigarette and using drugs, and feel safe to study at night in classrooms appeared as significant predictors of retention. The graduation rate and retention rate for the students who assigned to the field they did not choose were lower than that for those assigned to the field they chose. Those with first year CGPA less than 2.0 were having lower rates of graduation and retention than those having greater than 2.0. The graduation and retention rates for the students having higher preparatory average result and higher entrance exam score were higher than that for those having lower. The students having parents' income less than 500 were less likely to retain than those having parents' income greater than 1500. The retention rate for the students who were not satisfied with their instructors was lower than those were satisfied. The students who cannot organize their study and leisure time easily were less likely to retain than those can organize. In conclusion, the factors those mainly affect female students' graduation and retention were more of academic variables; hence we recommend that assigning to the field they choose by their interest may help female students' graduation and retention. The teaching method at secondary and preparatory schools should be designed to challenge and motivate them to adequately prepare them for Higher Education Institutions. Moreover, campus and Department administrators in collaboration with the students themselves and academic staff need to work hard to bring change in behavior, academics, and social aspects of female students at the University.

Keywords: Bayesian Logistic Regression, Stratified Sampling, Retention, Graduation, Female Students

1. Introduction

Education is the only essential measure that can guarantee women for their full participation in different development programs like leadership, health, education, Agriculture, nutrition, and socio-economic programs. Educated Women contribute to the society by improving quality of life and by enhancing national development through increasing economic production rate, better sanitation and nutritional practices, reduced child and maternal mortality rate. They are also mothers of better educated children. Women's participation in all fields of the world has become significant [1, 2].

Retention is the current enrolment in a program, which is the withholding of students in the program and is important for both the students and the University in the sense that the

students can complete the degree that they are striving for and the University is able to complete the goal of retaining the student [3]. Female Student's retention is one of the major concerns currently in higher education [4]. Many researchers found that retention rates for women in the higher education system are significantly lower than for men [5-9].

Graduation is the completion of a program and graduation rate is the completion rate [10]. If a female student retains all the consecutive semesters with higher performance, then we can be sure that she will graduate with in the latest year.

Increasing the retention and graduation rates of female students is vital to the mission of the University system [11]. The participation gap of females in education is wide in developing countries like Ethiopia, which has low females participation than their male counterparts [12]. The ministry of education in Ethiopia is taking many measures to improve girls' enrollment in the higher education. One of the measures is taking affirmative action strategies. As per the strategy, for Higher education enrollment, Ethiopian Higher Education Entrance Qualification Certificate (EHEEQC) grades favor girls in which for girls, the entrance exam score is lower than their male counterparts.

Women success can be measured based on the percentage of women graduates. Female graduates contribute more likely to avoid poverty and more likely to continue with graduate and PhD programs. They have also less economic risk than those dropouts [13]. Increasing the number of female graduates will not only benefit the individual but also will have broader public benefit such as paying tax.

Seyoum [14] indicated that the number of female graduates in Ethiopia between 1963 and 1973 years was only 6.7%. In Hawassa University (HU), the percentage of women graduates over the 1999/ 2001, 2001/ 02, 2002/ 03, 2003/ 04, and 2004/05 years did not show any meaningful change in terms of proportions. Out of the 27,209 enrolled between 2000/01 and 2004/05 academic years at undergraduate programs, the average proportion of women graduates over the same period was 11.7%, which is very low [15]. Furthermore; on the same area, on average about 1,451 female students in the Agriculture, Natural Science and Social Science faculties were enrolled in 2003/04 and 2004/05 academic years in which about 74% (1075) of them has retained [16].

The dropout rates and educational experiences of female students cannot be ignored [13]. A high rate of Female students' dismissal indicates a failure on the part of an institution to achieve its goal [17]. Not retaining has a negative psychological implication for the female students such as feelings of low self-esteem [4]. The many problems that drop out female students face lead them to abnormal socio-economic crisis [13].

Different questions arise as to whether female students' retention or graduation is the result of characteristics of individual students or of factors inherent in the structure, process and culture of education. Many studies have required identifying models and sets of variables to explain what affects female student's retention and graduation in the

Higher education system [16, 18-19].

1.1. Statement of the Problem

Female students are failing to graduate at alarmingly high rates [13]. The dropout rates of female students result in severe economic risks for themselves, their families, and the whole society. These dropout female students are not only at economic risk but also they are more likely to be unemployed, get lower salaries, and suffer from health problems [20].

The female students that do not retain, even though they are given a chance to return to the University after having one year break at their home do not want to go to their home. They are more likely to become psychologically ill, financially dependent, pregnant, single mothers, prostitutes, and street girls. If they even give birth, their children would also likely to drop out [13].

There are so many factors that may influence the retention and graduation of female students in Higher Education. Therefore the following research questions were countered in this study:

1) What are the retention and graduation rates of female students?
2) What are the factors that mainly affect the retention and graduation of female students?
3) Do female students in different fields of study have the same rates of retention and graduation?
4) Does the first year performance matter for graduation/ retention?

1.2. Objective of the Study

The general objective of this study was to identify the factors that influence the retention and graduation of female students at Higher Education Institution at HU, Ethiopia.

The specific objectives of this study are:-

1) To estimate the retention and graduation rates of female students.
2) To identify the major factors that influences the retention and graduation of female students.
3) To provide information for policy makers, University administrators, and researchers.

1.3. Significance of the Study

The study of analysis of Female students Retention and Graduation can be an example in the Higher Education Institution to suggest approaches for improving female student's success in different fields and also to help them to improve their academic competence by giving them information about how they can retain and graduate. It will be useful for the University by providing information about the factors that make female students to retain and graduate, for University administrators and researchers to give guidance to work more on this area and seek a better understanding of why female students are dismissed and also assist the administration to develop counseling and advising. It is also useful in providing the society information about the factors that may endanger the female students of higher Education.

2. Materials and Methods

2.1. Description of the Study Area and Population

The study was carried out at HU, which is located in the Southern Nations Nationalities People's Region (SNNPR). SNNPR is located in the South and South western part of Ethiopia. Hawassa is the regional capital city of SNNPR, which is one of the fastest growing and most dynamic cities in the country with high potential to attract researchers, residents, and investors. It is located in the Sidama Zone and is 275 kilometers south of Addis Ababa via Debre Zeit with a population of 159,013 [21]. HU was established in April 2000 and today it is a comprehensive University engaged in the provision of all-round education, research, training, and community service through its diversified areas of academic units. During the study period, the University had four campuses called Hawassa Main Campus, Hawassa College of Agriculture, Hawassa College of Medicine, and Wondogenet college of Forestry Campuses. The University has about 50 programmes of studies.

The study population included female students who were enrolled to the University in 2005/ 06, 2006/ 07, 2007/ 08, and 2008/ 09 academic years. The total size of the population was 1931, among which 415 enrolled in 2005/ 06, 358 in 2006/ 07, 331 in 2007/ 08, and 827 in 2008/ 09.

2.2. Data

The data for this study was obtained from Primary and Secondary data. The secondary data was collected from the HU registrar offices to identify factors that may predict female students' graduation. These data included data of 2005/06, 2006/07 and 2007/08 entries. The primary data was collected from 2008/09entry female students in HU Main Campus and College of Agriculture through questionnaire to identify the factors those may affect female student's retention. The data was collected by 8 trained enumerators and the researcher. The data satisfied the following criteria.

Exclusion and Inclusion Criteria: This study excluded female students who were having four year or above program of study. It also excluded female students who were before 2005/ 06and after 2008/ 09entries.

2.3. Sampling Design and Procedure

Stratified random sampling method was used as asampling design for selecting a representative sample of female students. It is a technique, which is explained by Cochran [22] and Al-Subaihi [23].

The stratification was based on 3 Faculties and 4 years of entry, which wasmade as follows:

Stratum 1: Female students who were 2005 entries in Agriculture with population size N_1 and sample size n_1.

Stratum 2: Female students who were 2006 entries in Agriculture with population size N_2 and sample size n_2.

Stratum 3: Female students who were 2007 entries in Agriculture with population size N_3 and sample size n_3.

Stratum 4: Female students who were 2008 entries in Agriculture with population size N_4 and sample size n_4.

Stratum 5: Female students who were 2005 entries in Natural Science with population size N_5 and sample size n_5.

Stratum 6: Female students who were 2006 entries in Natural Science with population size N_6 and sample size n_6.

Stratum 7: Female students who were 2007 entries in Natural Science with population size N_7 and sample size n_7.

Stratum 8: Female students who were 2008 entries in Natural Science with population size N_8 and sample size n_8.

Stratum 9: Female students who were 2005 entries in Social Science with population size N_9 and sample size n_9.

Stratum 10: Female students who were 2006 entries in Social Science with population size N_{10} and sample size n_{10}.

Stratum 11: Female students who were 2007 entries in Social Science with population size N_{11} and sample size n_{11}.

Stratum 12: Female students who were 2008 entries in Social Science with population size N_{12} and sample size n_{12}.

With the total number of female students of the study:

$$N = N_1+N_2+N_3+N_4+N_5+N_6+N_7+N_8+N_9+N_{10}+N_{11}+N_{12}$$

The total sample size of female students is:

$$N = n_1+n_2+n_3+n_4+n_5+n_6+n_7+n_8+n_9+n_{10}+n_{11}+n_{12}$$

2.3.1. Sample Size Determination

When Stratified random sampling technique is used to estimate the population proportion, the following formula for sample size n is used [22-23]. The formula that gives the sample size needed was:

$$n = \frac{\sum_{i=1}^{L}\left(\frac{N_i^2 p_i q_i}{W_i}\right)}{\frac{N^2 d^2}{Z^2} + \sum_{i=1}^{L} N_i p_i q_i} \qquad (1)$$

where

n = is the sample size needed, p_i = is the subpopulation proportion for stratum i, which is the probability that a female student retained or graduate. It is obtained from a previous study, which was conducted by Tesfaye [16]. Hence, for both retention and graduation, the estimate for p_i was taken to be 0.74, then $q_i = 1-p_i = 1-0.74 = 0.26$.

d = is the precision level, which is the margin of error. The specification of d must be small to have a good precision. In this study d = 0.03 was used to minimize cost.

L = 12 is the total number of strata, N_i= Total number of population in stratum i, W_i= is the estimated proportion of N_i to N, and $Z_{\alpha/2}$ = is the inverse of the standard normal cumulative distribution that correspond to the level of confidence, which is equal to the upper $\alpha/ 2$ point of standard normal distribution, where $\alpha = 0.05$, i.e., $Z_{\alpha/2} = 1.96$.

$N_1= 120$, $N_2 = 156$, $N_3 = 105$, $N_4 = 167$, $N_5 = 138$, $N_6 = 82$

$N_7 = 119$, $N_8 = 417$, $N_9 = 157$, $N_{10} = 120$, $N_{11} = 107$

$N_{12} = 243$, then, N = 1931

Using the formula in "(1)", the calculated sample size was to be n = 576. To compensate a non-response, 10% of 576 = 57.6 ≈ 58 was added to the computed n. Thus, the required sample size for this study was n = 634. Applying the strata weights, the sample allocation to each stratum was proportional to the total number of units in the stratum. The sample size for the ith stratumwas $n_i = W_i*n$, so that.

$$n_1 = 39, n_2 = 51, n_3 = 35, n_4 = 55, n_5 = 45, n_6 = 27, n_7 = 39$$

$$n_8 = 137, n_9 = 52, n_{10} = 39, n_{11} = 35, n_{12} = 80$$

Further, using proportional allocation of the above each sample sizes from each stratum, the sample size breakdown for each department in each respective year was also made.

2.4. Variables in the Study

Variables in this study were selected based on some past studies and thosewere expected to be factors that may affect female students' retention and graduation at HU. Some of these variables are continuous and others are categorical.

2.4.1. Graduation Variables

The dependent variable is female students' graduation status. It has two outcomes, graduated or not graduated and it is dichotomized as 1 if the student graduated and 0 if she did not (See Table 1).

Table 1. *List of variables and their codes in graduation analysis.*

Independent Variables	Value Labels
Year of entry	(0) 1998 entry, (1) 1999 entry, (2) 2000 entry
Age at first year	Interval
Marital Status	(0) Single, (1) Married
Region	(0) Gambela, (1) Afar, (2) BenshangulGumez, (3) Somalia, (4) Hareri, (5) Diredawa, (6) Amhara, (7) Tigray, (8) SNNPR, (9) Oromia, (10) Addis Ababa
Preparatory Average Result	Interval
Entrance Exam Score	Interval
Faculty	(0) Natural Science, (1) Social Science, (2) Agriculture
Choice of field	(0) No, (1) Yes by interest
First year CGPA	(0) < 2.0, (1) ≥ 2.0

2.4.2. Retention Variables

The dependent variable is female students' retention status. It has also two outcomes, retained coded as 1 or not retained coded as 0. They are classified as demographic and socio-economic variables, family backgrounds etc.

Table 2. *List of variables and their codes in retention analysis.*

Independent Variables	Value Labels
Age at the time of first year	Interval
Marital Status	(0) Single, (1) Married
Region	(0) Gambela, (1) Afar, (2) BenshangulGumez, (3) Somalia, (4) Hareri, (5) Diredawa, (6) Amhara, (7) Tigray, (8) SNNPR, (9) Oromia, (10) Addis Ababa
Preparatory Average Result	Interval
Entrance Exam Score	Interval
Faculty	(0) Natural Science, (1) Social Science, (2) Agriculture
Choice of field	(0) No, (1) Yes by interest
First Year CGPA	(0) < 2.0, (1) ≥ 2.0
Religion	(0) Others, (1) Muslim, (2) Orthodox, (3) Protestant
Mother"s Education	(0) Illiterate, (1) Informal Education, (2) Grade 1-8, (3) Grade 9-12, (4) College certificate/ diploma, (5) Degree and above
Father"s Education	(0) Illiterate, (1) Informal Education, (2) Grade 1-8, (3) Grade 9-12, (4) College certificate/ diploma, (5) Degree and above
Mother"s occupation	(0) No job, (1) Farmer, (2) Run own private business, (3) Employed
Father"s occupation	(0) No job, (1) Farmer, (2) Run own private business, (3) Employed
Parents Income	(0) < 500, (1) 500-1000, (2) 1001-1500, (3) > 1500
Residence before joining University	(0) Rural, (1) Urban
Place of Food Service	(0) None Café, (1) In Café
Pregnancy	(0) No, (1) Yes
Habit of chewing chat	(0) No, (1) Yes
Habit of drinking alcohol	(0) No, (1) Yes
Habit of smoking cigarette and using drugs	(0) No, (1) Yes
Satisfaction on Dormitory	(0) Low, (1) Medium, (2) High
Satisfaction on laboratory*	(0) D, (1) ID, (2) A

* D = Disagree, ID = Indifferent, A = Agree.

2.5. Statistical Models

In this study, Descriptive Statistics and Bayesian logistic regression was used. Logistic regression was used to test for each predictor's significance while controlling other predictors.

2.5.1. Bayesian Logistic Regression

Statistical inferences are usually based on maximum likelihood estimation (MLE), which chooses the parameters that maximize the likelihood of the data. In MLE, parameters are assumed to be unknown but fixed, and are estimated with some confidence. In Bayesian analysis, the uncertainty about the unknown parameters is quantified using probability so that the unknown parameters are regarded as random variables.

Bayesian inference is the process of analyzing statistical

models with the incorporation of prior knowledge about the model or model parameters. The source of such inference is

Thomas Bayes. This is done by applying Bayes' theorem. The posterior distribution is written as:

$$\text{Pr}(parameters \mid data = \frac{\text{Pr}(data \mid parameters) \times \text{Pr}(parameters)}{\text{Pr}(data)} \propto Likelihood \times \text{Pr}ior \qquad (2)$$

The prior distribution expresses the information available to the researcher before any "data" are involved in the statistical analysis.

Logistic regression is appropriate for data where the response variable is dichotomous and Bayesian logistic regression procedure is used to make inference about the parameters of a logistic regression model. Bayesian inference for logistic Regression analyses follows the usual pattern for all Bayesian analysis. The basic steps and concepts that should be considered in analyzing Bayesian inference should be the likelihood function of the data, a prior distribution over all unknown parameters, and the posterior distribution over all parameters.

2.5.2. Likelihood Function

The joint distribution of n independent Bernoulli trials is the product of each Bernoulli distributions and the sum of independent and identically distributed Bernoulli trials makes up a Binomial distribution.

Let Y_1, Y_2,..., Y_n be independent Bernoulli trials having probability of success p_1, p_2, ..., p_n respectively. That is, $y_i = 1$ with probability of p_i or $y_i = 0$ with probability of $1 - p_i$, for i = 1, 2,..., n. Since the trials are independent, the joint distribution of Y_1, Y_2, ..., Y_n is the product of n Bernoulli probabilities. That is:

$$L(\text{data}|\beta) = \prod_{i=1}^{n} P(Y_i \mid X_{i1}, X_{i2}, \dots X_{im}) = \prod_{i=1}^{n} p_i{}^{Y}[1-p_i]^{1-Y} \quad (3)$$

where p_i represents the probability of retaining or graduating for female student i with covariate vector X_i, $Y_i = 1$ represents for retained or graduated, and $Y_i = 0$ represents for did not

retain or did not graduate.

2.5.3. Prior Distribution

It is important to specify the prior distribution in Bayesian inference because it influences the posterior inference. Emphasis has to be given in specifying the prior mean and variance. The prior mean provides a prior point estimate for the parameter of interest, while the variance expresses the uncertainty concerning this estimate.

In general, any prior distributions can be used, depending on the available prior information. When there is some information about the likely values of the unknown parametersβ_0, β_1, β_2, ..., β_m, then informative priors are used but when no prior information is available, priors, which are called non informative or vague priors are used.

Most of the time, priors with mean zero and large variance are most common priors for logistic regression parameters. The assumed prior normal distribution for parameter is given by:

$$f(\beta_j) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\beta_j - \mu_j}{\sigma_j}\right)^2\right) \qquad (4)$$

where μ_j is the normal distribution mean and σ_j is the variance.

2.5.4. The Posterior Distribution

The posterior distribution is obtained from the product of the full likelihood function and the prior distribution over all parameters. The posterior distribution is written as:

$$f(\beta \mid \text{data}) \propto \prod_{i=1}^{n} p_i{}^{Y}[1-p_i]^{1-Y} \times \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\beta_j - \mu_j}{\sigma_j}\right)^2\right) \qquad (5)$$

wheref (β| data) is the posterior distribution, which is the product of the logistic regression likelihood and the normal prior distributions for theβparameters.

The mean of the posterior distribution can be used as a point estimate of βbut computing the estimate of $\beta$from the posterior distribution may be difficult analytically. Therefore non – analytical methods like simulation techniques are used. Of these techniques Markov Chain Monte Carlo (MCMC) methods is the most popular. Hence, the posterior mean of the parameters' can be obtained from the mean of the sampled values of the Markov Chain after the burn-in period.

2.5.5. Assessing the Bayesian Logistic Regression Model

i. Markov Chain Monte Carlo Methods (MCMC):

MCMC methods are attempted to simulate direct draws from some complex distribution of interest. The most widely

used MCMC technique is the Gibbs sampler. The MCMC methods are explained in detail by Stamps [24].

ii. Convergence of the Algorithm:

Before simulated parameters are summarized, we must ensure that the chains have converged. Several diagnostic tests were developed to monitor the convergence of the algorithm. Among several ways, the most popular and straight forward convergence assessment methods are: Autocorrelation, Time series plots, Gelman-Rubin statistic, and Density plot [25].

iii. The Burn-in Period:

Burn-in is aninformal term that describes the practice of throwing away some iteration at the beginning of an MCMC run. It is explained by Merkle et al. [26].

iv. Assessing Accuracy of the Bayesian Logistic Regression:

Once we achieved convergence, we need to run the

simulation for a further number of iterations to obtain samples that can be used for posterior inference. The simplest way to assess the accuracy of the posterior estimates is by calculating the Monte Carlo error (MCE) for each parameter, indicating the quantity of interest is calculated with precision [24].

3. Results and Discussion

The sample size determined for this study was 634. However there were 29 non-respondents and so the data analyzed in this study were based on 605 respondents. All computations were conducted in Statistical Package for Social Science (SPSS) version 13.0 and WinBUGS version 14.0.

The analysis has two parts: one for graduation and the other for retention. In the first case, 362 female students enrolled in the academic years 2005/ 06-2007/ 08 with 37.6% from 2005/ 06, 32.3% from 2006/ 07, and 30.1% from 2007/08 entries were considered. In terms of their fields of study 30.7% were from Natural Science, 34.5% were from Agriculture, and 34.8% were from Social Science.

For the second case, 243 female students who enrolled to the University in 2008/09 academic year were considered. 47.7% were from the Natural Science, 19.3% were from Agriculture, and 32.9% were from Social Science.

3.1. Results from Analysis of Graduation Data

From Table3, among the 362 female students of 2005/06-2007/08 entries, 72.1% graduated and the rest 27.9% did not. Regarding their background, about 72.4% were assigned to the field of their choice of interest while 27.6% of them were not assigned to the field they chose. During their first year stay, 58.3% of them scored CGPA of 2.00 and above while the rest 41.7% scored CGPA of less than 2.00. The mean of their first year CGPA was 2.10 with standard deviation 0.7198. The mean entrance exam score of the respondents was 238.31 with standard deviation of 35.097, while that of preparatory average result was 69.07 with standard deviation of 6.2839. The mean age of the respondents was 19.09 with standard deviation 1.048. 99.7% were never married.

Table 3. *Descriptive Statistics of Continuous Predictor Variables.*

Variables	N	Min	Max	Mean	Std. Error	S. D.
Age at first year	362	16	24	19.09	0.055	1.048
Entrance Exam Score	362	158	359	238.31	1.845	35.097
First Year CGPA	362	0.28	3.89	2.1015	0.0378	0.7198
Preparatory Average Result	362	52.00	87.15	69.07	0.3303	6.2839

3.1.1. Univariate Analysis Results

Table 4 displays the Wald test for the continuous covariates age, entrance exam score, and preparatory average result. The tests are all significant at 5% level.

Table 4. *Univariate Logistic Regression Results for Continuous Covariates.*

Variables	Wald statistics	d. f.	Sig.
Age at first year	64.428	1	0.000*
Entrance exam score	79.518	1	0.000*
Preparatory average result	76.010	1	0.000*

3.1.2. Bivariate Analysis Results

As can be seen in Table 5, the graduation rate was estimated to be 84.7% among the respondents who enrolled to the field of their choices. The estimate was 76.6% for Natural Science, 71.2% for Agriculture, and 69.0% for Social Science. The graduation rate among those having first year CGPA greater or equals to 2.00 was 93.8%. Per entry, it was 71.3% for 2005/06, 75.2% for 2006/07, and 69.7% for 2007/ 08. For the chi-square tests, it was found that graduation status of the female students considered was associated with their field choices and first year CGPA.

Table 5. *Association between female students' graduation status and explanatory variables (HU, 2005/ 06– 2006/ 07 Entry).*

Variables	N	%	Grad. (%)	Not Grad. (%)	d. f.	Chi- Square	LR
Graduation Status	362	100	72.1	27.9			
Marital Status							
Single	361	99.7	72.0	28.0			
Married	1	0.3	100	00.0			
Choice of Field							
No	100	27.6	39.0	61.0	1	75.25	70.95
Yes	262	72.4	84.7	15.3			
Sig. (2-sided)						0.000*	0.000*
First Year CGPA							
< 2.0	151	41.7	41.7	58.3	1	118.84	15.80
≥ 2.0	211	58.3	93.8	6.2			
Sig. (2-sided)						0.000*	0.000*
Year of Entry							
2005 entry	136	37.6	71.3	28.7			
2006 entry	117	32.3	75.2	24.8	2	0.910	0.919
2007 entry	109	30.1	69.7	30.3			
Sig. (2-sided)						0.634	0.632
Faculty							

Variables	N	%	Grad. (%)	Not Grad. (%)	d. f.	Chi- Square	LR
Natural Sci.	111	30.7	76.6	23.4			
Social Sci.	126	34.8	69.0	31.0	2	1.740	1.767
Agriculture	125	34.5	71.2	28.8			
Sig. (2-sided)						0.419	0.413
Region							
Amhara	61	16.9	70.5	29.5			
Tigray	20	5.5	75.0	25.0			
SNNPR	125	34.5	72.0	28.0	5	1.450	1.428
Oromia	74	20.4	71.6	28.4			
Addis Ababa	63	17.4	76.2	23.8			
Others	19	5.2	63.2	36.8			
Sig. (2-sided)						0.919	0.921

*Significant at 5% level.

3.1.3. Bayesian Logistic Regression Results for Graduation Data

Based on the Wald statistic and chi – square test results, 5 covariates with P_value less than 0.05 were selected for the Bayesian logistic regression analysis. These were choice of field (b. CH), first year CGPA (b. FGP), entrance exam score (b. EN), preparatory average result (b. EN), and age at first year (b. A) variables.

Three parameter chains were set up to be sampled for 300000 iterations each. The first 30000 iterations were discarded from each chain, leaving a total sample of 810000 to summarize. Four of the predictor variables except age at first year (b. A) were significant. This is because the 95.5% confidence interval for b. A contains zero (See in Table 6).

Table 6. *Summary Statistics of the Posterior Distribution of the Model Parameters Obtained from the Sampled Values.*

Variables	Mean ($\hat{\beta}$)	sd	MC error	0.05*sd	95%CI Lower	Upper
alpha	-22.7	4.413	0.0956	0.221	-31.61	-14.19*
b. A	0.136	0.1556	0.0031	0.008	-0.168	0.4417
b. CH	1.448	0.3804	0.0012	0.019	0.709	2.201*
b. EN	0.030	0.0068	0.0001	0.000	0.018	0.044*
b. FGP	2.333	0.4091	0.0022	0.021	1.554	3.159*
b. PAR	0.181	0.0373	0.0007	0.002	0.108	0.255*

i. Assessing Convergence for Graduation Data.

In the absence of conjugate prior distributions, WinBUGS uses sampling methods. These sampling methods typically guarantee that, under regularity conditions, the resulting sample converges to the posterior distribution of interest. Thus, before we summarize simulated parameters, we must ensure that the chains have converged.

a. Time Series Plots:

If the plot looks like a horizontal band, with no long upward or downward trends, then we have evidence that the chain has converged. Here, the three independently generated chains mixed together or overlapped and these plots display good mixing of the chains.
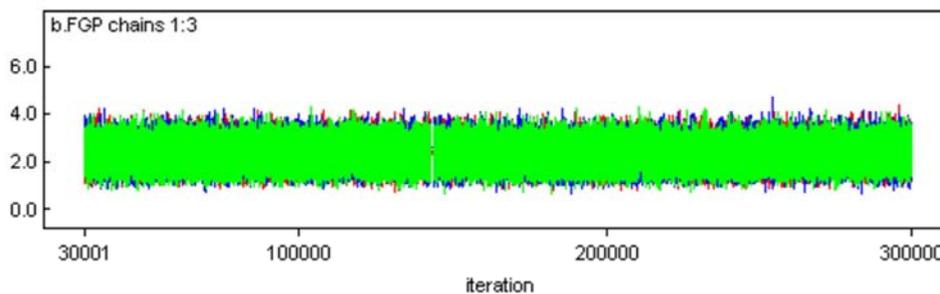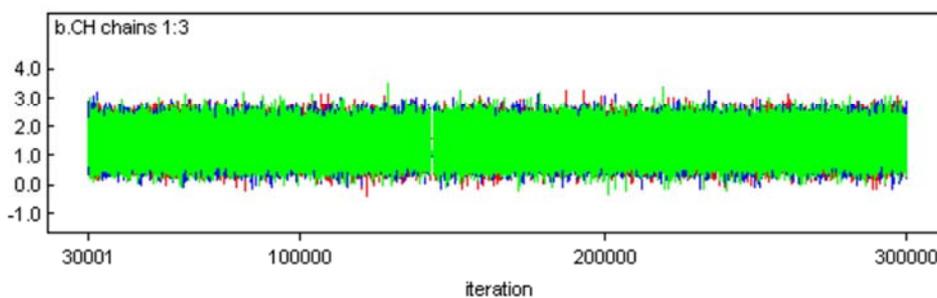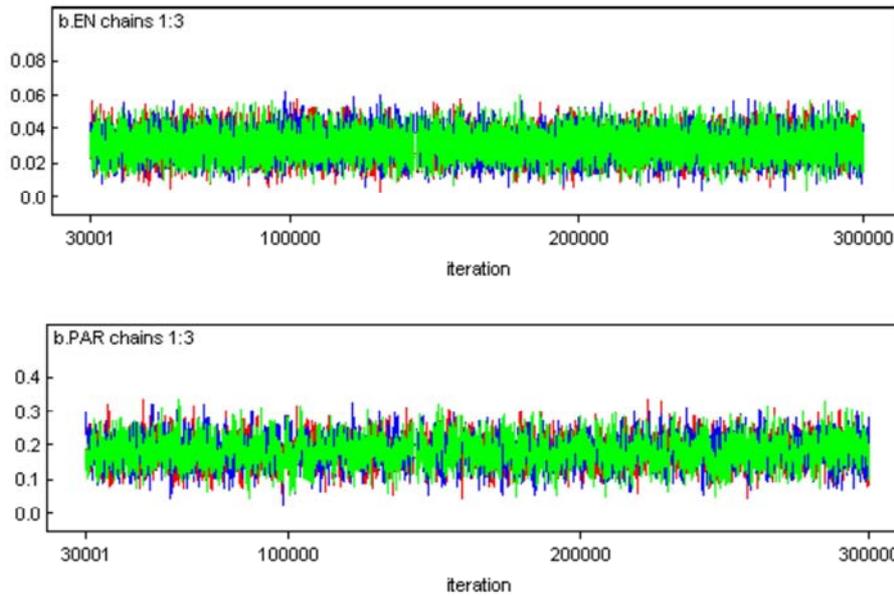
Figure 1. *Convergence of Time Series Plots for the Coefficients of Choice of Field (b. CH), first year CGPA (b. FGP), entrance exam score (b. EN), and preparatory average result (b. EN) in the Graduation Analysis Data.*

b. Autocorrelation Plot:

Autocorrelation plots can also be used to test for convergence. High autocorrelations in parameter chains show that a model that is slow to converge. Here, the plots showed that the three independent chains were mixed or overlapped to each other and died out for higher lags and hence this is an evidence of convergence (Results are displayed in Figure 2).
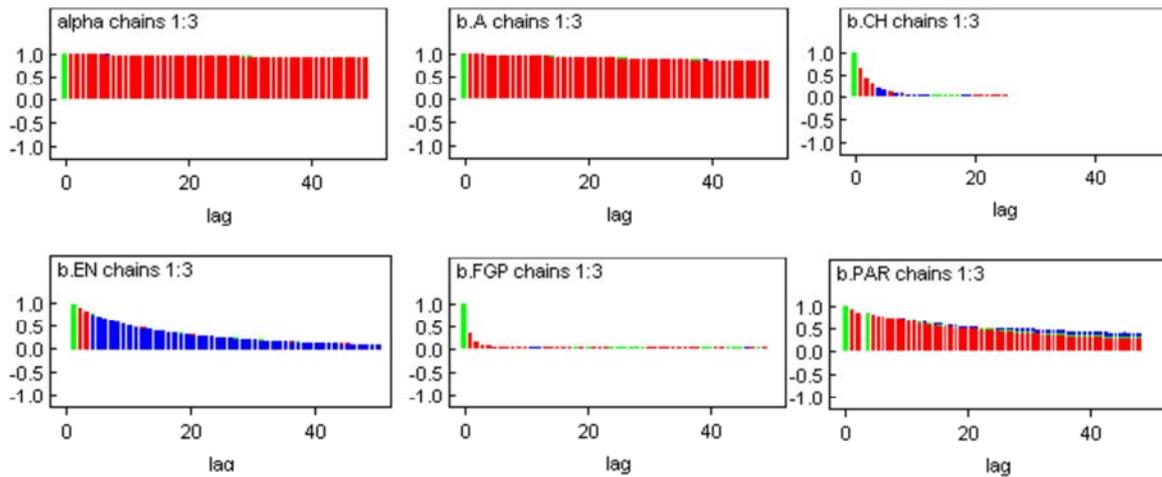

Figure 2. *Convergence assessment using autocorrelation plot for the coefficients of the covariates in Graduation Analysis.*

c. Gelman-Rubin Statistic (GR):

For a given parameter, this statistic assesses the variability within parallel chains as compared to variability between parallel chains. The model is judged to have converged if the ratio of between to within variability is close to 1. From the Plots of this statistic, the green line represents the between variability, the blue line represents the within variability, and the red line represents the ratio. Evidence for convergence comes from the red line being close to 1 on the y-axis and from the blue and green lines being stable across the width of the plot [26]. Hence, evidence for convergence has been reached (Results are displayed in Figure 3).
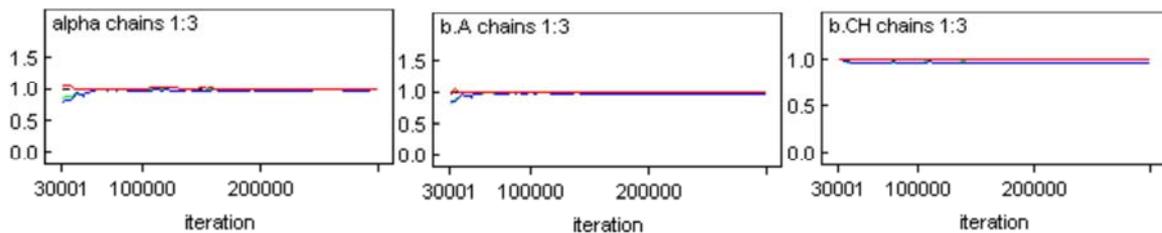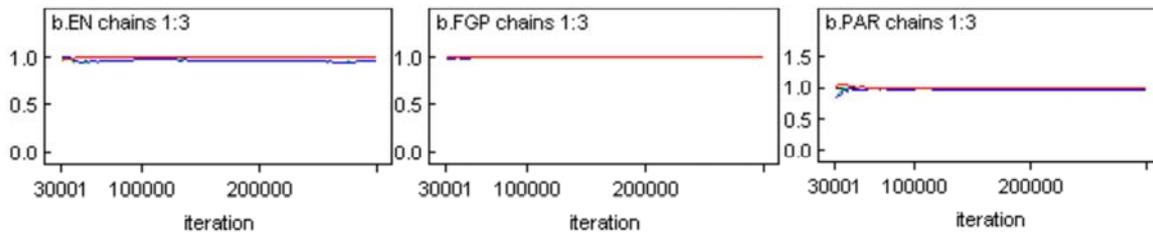
***Figure 3.** ConvergenceusingGelman-Rubin Statistic for the parameters of Choice of Field (b. CH) and First Year CGPA (b. FGP) in the Graduation Analysis Data.*

d. Density plot:

This is another recommended technique for identifying convergence. It is a smoothed kernel density estimate for continuous variable or a histogram for discrete variable. Since

the plots for the most predictor variables indicated that the coefficient has normal distribution, the simulated parameter value indicated convergence (See Figure 4).
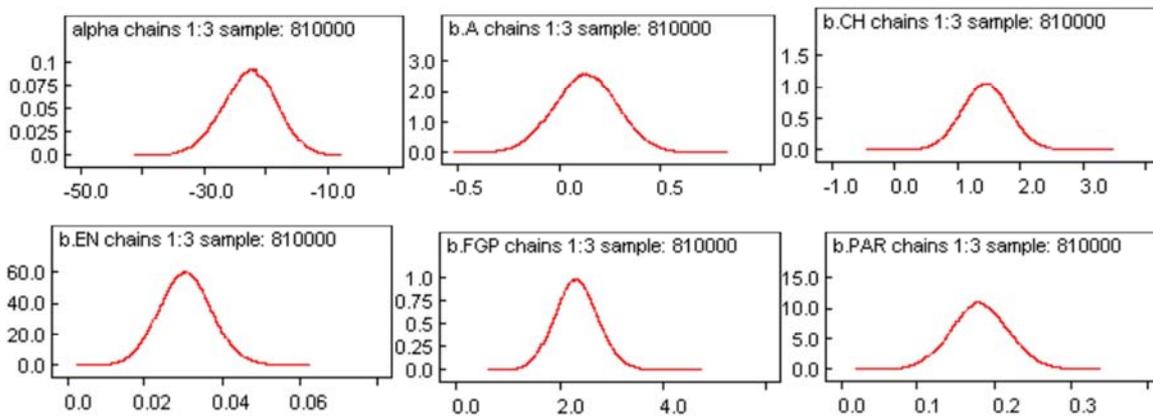

***Figure 4.** Density Plots of the Posterior Distributions of the Parameters in Graduation Analysis.*

ii. Assessing Accuracy of Bayesian Logistic Regression Model of Graduation Data.

Once convergence has been achieved, we need further simulation for a further number of iterations to obtain samples that can be used for posterior inference. The more samples we have, the more accurate the posterior estimates will be. One way to assess the accuracy of the posterior estimates is by calculating the Monte Carlo error for each parameter. This is an estimate of the difference between the mean of the sampled values and the true posterior mean. As a rule of thumb, the simulation should be run until the Monte Carlo error for each parameter of interest is less than about 5% of the sample standard deviation.

Table 7 contains the estimated coefficients, Mean, the standard deviation (sd), Monte Carlo (MC) errors, 5% of the standard deviation (0.05*sd) and 95.5% Confidence Interval. The MC error for choice of field (CH) is 0.0012 and this is less than 0.05*sd = 0.01902. MC error for entrance exam score is 0.0001 and this is less than 0.05*sd = 0.00034, etc. This implied that the MC error for each significant predictor variable is less than 5% of its posterior standard deviation and hence

convergence and accuracy of posterior estimates are attained and the model is appropriate to estimate posterior statistics.

Confidence intervals for the constant (alpha) and for the predictor variables choice of field (b. CH), entrance exam score (b. EN), first year CGPA (b. FGP) and preparatory average result (b. PAR) do not include zero. Therefore, these predictor variables were significant and hence mainly affect female student's graduation. The Mean values for b. CH, b. EN, b. FGP, and b. PAR predictor variables were 1.448, 0.0304, 2.333, and 0.1806 respectively. These values are all positive indicating positive relationship between the dependent variable graduation and these predicator variables. For example, in the case of choice of field (b. CH), since 'No' is coded as 0 and 'Yes' is coded as 1, the positive value indicated that female students who assigned to the field they chose were more likely to graduate than those who didn't assigned to their field choice by their interest. In the case of entrance exam score (b. EN), female students having higher entrance exam score were more likely to graduate than those having lower entrance exam score (Results are displayed in Table 7).

***Table 7.** Summary Statistics of the Posterior Distribution of the Model Parameters Obtained from the Sampled Values (HU, 2005/ 06 – 2006/ 07Entry).*

Node	Mean ($\hat{\beta}$)	sd	MC error	0.05*sd	95.5% C. I.	
					Lower	Upper
alpha	-22.69	4.413	0.0956	0.22065	-31.61	-14.19*
b. A	0.1355	0.156	0.0031	0.00778	-0.1681	0.4417
b. CH	1.448	0.380	0.0012	0.01902	0.7086	2.201*
b. EN	0.0304	0.007	0.0001	0.00034	0.0175	0.0442*
b. FGP	2.333	0.409	0.0022	0.02046	1.554	3.159*
b. PAR	0.1806	0.037	0.0007	0.00187	0.1084	0.2551*

3.1.4. Interpretation and Discussion for the Graduation Analysis Data Results

This study tried to estimate female students' graduation rate and also has provided some information about the factors that determine female students' graduation at Hawassa University. According to the results, among 362 female students of the academic year 2005/06-2007/08 with mean age at first year equals 19.09; 72.1% graduated but the rest 27.9% did not. About 37.6% were 2005/06 entry, 32.3% were 2006/07 entry and 30.1% were 2007/08 entry. 30.7% of them were from the Natural Science Faculty, 34.8% were from the Social Science Faculty and 34.5% were from the Agriculture Faculty. About 72.4% were assigned to the field they chose by their interest but 27.6% of them did not. During their first year stay, 58.3% of them scored CGPA of 2.00 and above while the rest 41.7% scored CGPA of less than 2.00.

From the bivariate and Univariate results, choice of field, first year CGPA, entrance exam score, preparatory average result and age at first year variables were found significant predicators of the dependent variable Graduation.

From the Bayesian logistic regression analysis, the constant (alpha) and the predictor variables choice of field (b. CH), entrance exam score (b. EN), first year CGPA (b. FGP) and preparatory average result (b. PAR) were significant predictors of female students' graduation. In this case all Mean ($\hat{\beta}$) values of the significant predictors were positive indicated positive relationship between the dependent variable graduation and these predicator variables. One variable that affects female students' graduation was preparatory average result (b. PAR). Female students having higher preparatory average result were more likely to graduate than those having lower preparatory average result and similar results were found by earlier studies conducted by Zhang et al. [11].

3.2. Results from Analysis of Retention Data

From the descriptive results in Table 8, among the 243 female students who were 2008/09 entry, 75.7% (184) retained but the rest 24.3% did not retain.

The mean age of the participants and its standard deviation were 18.576 and 0.7856 respectively, the mean entrance exam score of the students and its standard deviation were 241.90 and 32.429, the average first year CGPA of the students and its standard deviation were 2.352 and 0.4970 and the mean of preparatory average result of the students and its standard deviation were 71.116 and 6.2114 respectively (Results are displayed in Table 8).

Table 8. Descriptive Summary Statistics of Continuous Predictor Variables.

Variables	N	Min	Max	Mean	Std. Error	S. D.
Age at first year	243	17	21	18.576	0.0504	0.7856
Entrance Exam Score	243	175	358	241.90	2.080	32.429
First Year CGPA	243	1.03	3.84	2.352	0.0319	0.4970
Preparatory Average Result	243	54.50	92.05	71.116	0.3985	6.2114

3.2.1. Univariate Analysis Results for Retention Data

The results in Table 9 indicated that the Wald statistic for entrance exam score and preparatory average result predictor variables is significant at 5% level of significance. Hence, these predictor variables were selected for the logistic regression analysis.

Table 9. Univariate Logistic Regression Results for Continuous Covariates.

Explanatory variables	Wald statistic	d. f.	Sig.
Entrance Exam Score	14.317	1	0.000*
Preparatory Average Result	50.417	1	0.000*

3.2.2. Bivariate Analysis Results for Retention Data

About 95.9% (233) were single but 4.1% were married. About 20.2% were from the region Amhara, 17.3% were from the region SNNPR, 28.4% were from the region Oromia, 28.8% were from the region Addis Ababa, 2.9% were from the region Tigray and 2.5% were from other than these regions. In the case of choice of field, 'Yes' stands for the 1st or 2nd choices but 'No' stands for the student's 3rd or above choices. Hence 79.4% were assigned to their 1st or 2nd choices and the rest 20.6% were assigned to their 3rd or above choices. About 3.3% of female students were having a father with no job, 26.7% were having a father who is a farmer, 23.5% were having a father who runs his own business and 46.5% were having an employed father. The percent of female students having their parents' income per month less than 500, between 500 and 1000, between 1001 and 1500, and greater than 1500 was 22.6%, 17.7%, 19.8% and 39.9% respectively. The percent of female students who ever got pregnancy was 5.8%. The percent of female students who were having a habit of chewing Khat was 8.6% and that of having a habit of smoking cigarette and using drugs was 3.3%. During their first year stay, 77.8% of them were having a CGPA of greater than or equal to 2.0 but the rest 22.2% of them were having a CGPA of less than 2.0 (See Table 10).

From Table 11, about 53.5% of female students were satisfied with their instructors. About 77.0% of them were committed to their academic activities. 71.6% of them were able to organize their study and leisure time easily. 51.9% of female students were able to focus on exam. 40.7% of them were being nervous and forgot facts they know on exam. About 37.9% of female students felt that they never achieve academic respects even if they work hard. The sociability of 66.7% female students helped them to reach their educational goals. About 36.2% of them felt safe to study in classrooms at night. About 62.1% of them matched their expectation about the field with the reality they faced.

Among the 79.4% of female students who were assigned to their 1st or 2nd choices, 87.6% of them retained but among the

20.6% who were assigned to their 3rd or above choices only 30.0% retained. Among the 94.2% female students who did not get pregnancy 79.9% retained but among the 5.8% ever got pregnancy female students, only 7.1% retained. Among the 53.5% who were satisfied with their instructors, 80.8% retained but among the 25.9% who were not satisfied with their instructors 60.3% retained. Among the 77.0% who were committed on their academic activities, 80.2% retained. Among the 77.8% who have their first year CGPA greater or equal to 2.0, 86.8% retained but among the 22.2% who have first year CGPA less than 2.0, 37.0% retained

As we can see from Table 10 and Table 11, the asymptotical significance value for both chi – square and likelihood ratio tests and for both Kendall's tau-b and tau-c tests is below 0.05 for predictor variables: choice of field, first year CGPA,

pregnancy, parent income, habit of chewing Khat, habit of smoking cigarette and using drugs, focus on exam, commitment to academic activities, organize study and leisure time, satisfaction with instructors and never achieve academic respects. This indicated that there is a significant relationship between the dependent variable retention and these predictor variables. The following predictor variables are also significant at 25% level of significance. These are: father occupation, being nervous and forget facts on exam, help of sociability in campus to reach educational goals, feel safe to study at night in classrooms, expectation about field and weak because of personal problems and hence these predictor variables can be selected with the other significant predictor variables for the logistic regression analysis.

Table 10. *The Distribution of Female Students' Retention Status with Respect to the Categorical Explanatory Variables (HU, 2008/ 09 Entry).*

Variables	N	%	Ret.	Not Ret.	d. f	Chi – square	LR
Retention Status	243	100%	75.7%	24.3%			
Marital Status							
Single	233	95.9%	76.0	24.0%	1	0.186	0.177
Married	10	4.1%	70.0%	30.0%			
Sig. (2-sided)						0.667	0.674
Region							
Amhara	49	20.2%	71.4%	28.6%	5	3.773	4.217
Tigray	7	2.9%	100%	0.0%			
SNNPR	42	17.3%	69.0%	31.0%			
Oromia	69	28.4%	78.3%	21.7%			
Addis Ababa	70	28.8%	77.1%	22.9%			
Others	6	2.5%	83.3%	16.7%			
Sig. (2-sided)						0.438	0.377
Choice of Field							
No	50	20.6%	30.0%	70.0%	1	71.578	63.481
Yes	193	79.4%	87.6%	12.4%			
Sig. (2-sided)						0.000**	0.000**
Father Occupation							
No Job	8	3.3%	75.0%	25%	3	4.728	4.566
Farmer	65	26.7%	66.2%	33.8%			
Runs Business	57	23.5%	77.2%	22.8%			
Employed	113	46.5%	80.5%	19.5%			
Sig. (2-sided)						0.193*	0.207*
Parent Income							
< 500	55	22.6%	60.0%	40.0%	3	14.610	14.280
500 – 1000	43	17.7%	67.4%	32.6%			
1001 –1500	48	19.8%	83.3%	16.7%			
> 1500	97	39.9%	84.5%	15.5%			
Sig. (2-sided)						0.002**	0.003**
Pregnancy							
No	229	94.2%	79.9%	20.1%	1	38.002	32.439
Yes	14	5.8%	7.1%	92.9%			
Sig. (2-sided)						0.000**	0.000**
Habit of Chewing Chat							
No	222	91.4%	77.9%	22.1%	1	6.811	5.966
Yes	21	8.6%	52.4%	47.6%			
Sig. (2-sided)						0.009**	0.015**
Habit of smoking Cigarette and using Drugs							
No	235	96.7%	77.4%	22.6%	1	11.575	9.489
Yes	8	3.3%	25.0%	75.0%			
Sig. (2-sided)						0.001**	0.002**
First Year CGPA							
< 2.0	54	22.2%	37.0%	63.0%	1	56.510	50.513
≥ 2.0	189	77.8%	86.8%	13.2%			
Sig. (2-sided)							0.000**

**Significant at 5% level and *Significant at 25% level.

***Table 11.** The Distribution of Female Students' Retention Status with Respect to the Ordinal Explanatory Variables (HU, 2008/ 09 Entry).*

Variables	N	Total%	Retained%	Not-Retained%	Kendall"s tau-b	Kendall"s tau-c
Satisfaction with Instructors						
Disagree	63	25.9%	60.3 %	39.7%	0.164	0.154
Indifferent	50	20.6%	82.0 %	18.0%		
Agree	130	53.5%	80.8 %	19.2%		
Sig. (2-sided)					0.011**	0.011**
Committed to academic activities						
Disagree	19	7.8%	78.9%	21.1%	0.168	0.125
Indifferent	37	15.2%	51.4%	48.6%		
Agree	187	77.0%	80.2%	19.8%		
Sig. (2-sided)					0.014**	0.014**
Organize study and leisure time						
Disagree	41	16.9%	48.8%	51.2%	0.220	0.178
Indifferent	28	11.5%	82.1%	17.9%		
Agree	174	71.6%	81.0%	19.0%		
Sig. (2-sided)					0.002**	0.002**
Focus on Exam						
Disagree	77	31.7%	62.3%	37.7%	0.153	0.145
Indifferent	40	16.5%	87.5%	12.5%		
Agree	126	51.9%	80.2%	19.8%		
Sig. (2-sided)					0.018**	0.018**
Being nervous and forget facts on exam						
Disagree	116	47.7%	78.4%	21.6%	- 0.079	- 0.074
Indifferent	28	11.5%	82.1%	17.9%		
Agree	99	40.7%	70.7%	29.3%		
Sig. (2-sided)					0.208*	0.208*
Never achieve academic respects						
Disagree	124	51.0%	81.5%	18.5%	0.128	0.118
Indifferent	27	11.1%	70.4%	29.6%		
Agree	92	37.9%	69.6%	30.4%		
Sig. (2-sided)					0.039**	0.039**
Help of Sociability in campus to reach educational goals						
Disagree	34	14.0%	73.5%	26.5%	0.076	0.065
Indifferent	47	19.3%	68.1%	31.9%		
Agree	162	66.7%	78.4%	21.6%		
Sig. (2-sided)					0.227*	0.227*
Feel safe to study at night in Classrooms						
Disagree	106	43.6%	81.1%	18.9%	- 0.116	- 0.113
Indifferent	49	20.2%	75.5%	24.5%		
Agree	88	36.2%	69.3%	30.7%		
Sig. (2-sided)					0.056*	0.056*
Expectation about Field						
Disagree	58	23.9%	72.4%	27.6%	0.079	0070
Indifferent	34	14.0%	67.6%	32.4%		
Agree	151	62.1%	78.8%	21.2%		
Sig. (2-sided)					0.206*	0.206*
Weak because of personal problems						
Disagree	118	48.6%	78.8%	21.2%	- 0.106	- 0.101
Indifferent	36	14.8%	86.1%	13.9%		
Agree	89	36.6%	67.4%	32.6%		
Sig. (2-sided)					0.093*	0.093*

**Significant at 5% level and *Significant at 25% level.

3.2.3. Bayesian Logistic Regression Results for Retention Data

The predictor variables those were significant in the Bivariate and Univariate analyses were selected for the Bayesian logistic regression. These were choice of field, first year CGPA, entrance exam score, preparatory average result, parent income, habit of chewing Khat, habit of smoking cigarette and using drugs, pregnancy, focus on exam, commitment to academic activities, organize study and leisure time, satisfaction with instructors, father occupation, being nervous and forget facts on exam, help of sociability in campus to reach educational goals, never achieve academic respects, feel safe to study at night in classrooms, expectation about field, and weak because of personal problems.

Three parameter chains were set up to be sampled for 400000 iterations each. The first 150000 iterations were discarded from each chain, leaving a total sample of 750000 to summarize. 11 predictor variables were significant in the Bayesian logistic regression (See Table 12).

i. Assessing Convergence for Retention Data

Before we summarize simulated parameters, we must ensure that the chains have converged.

a. Time series plots for retention Data
The three independently generated chains mixed together or overlapped and these plots display good mixing of the chains (Results are displayed in Figure 5).


alpha chains 1:3

b.CH chains 1:3

b.Chew chains 1:3

b.Commt chains 1:3

b.EN chains 1:3

b.Exptfld chains 1:3
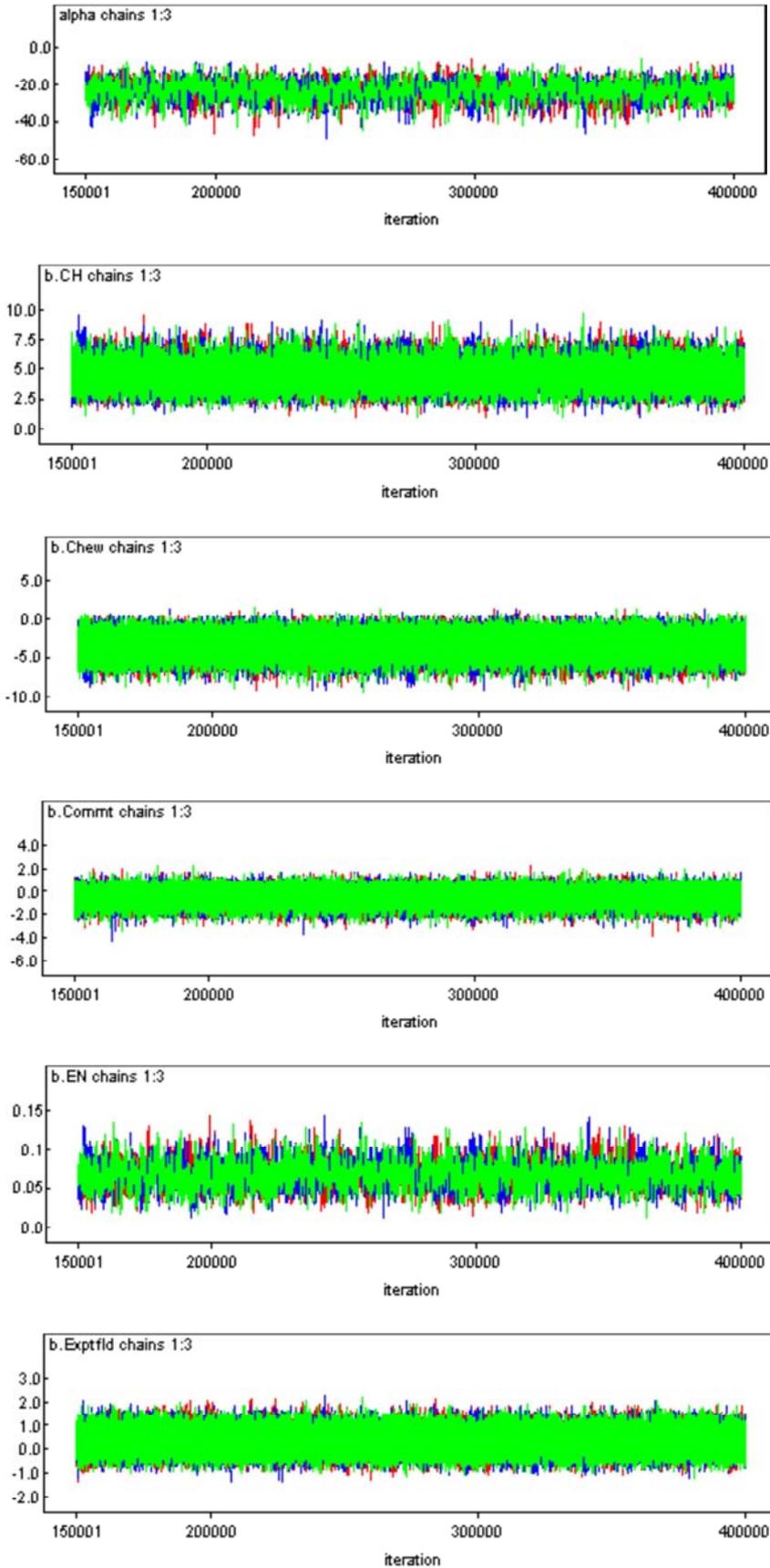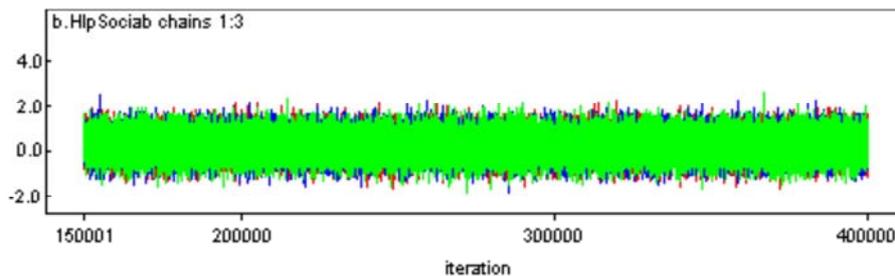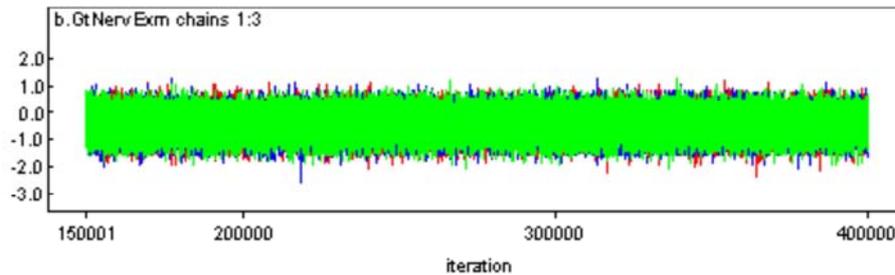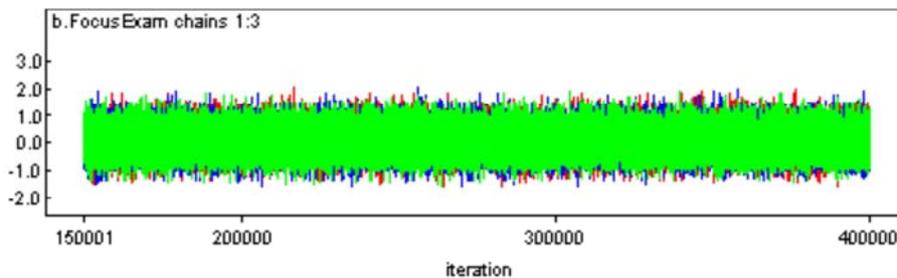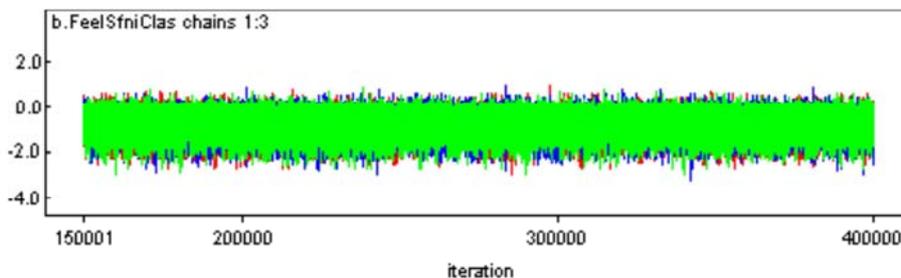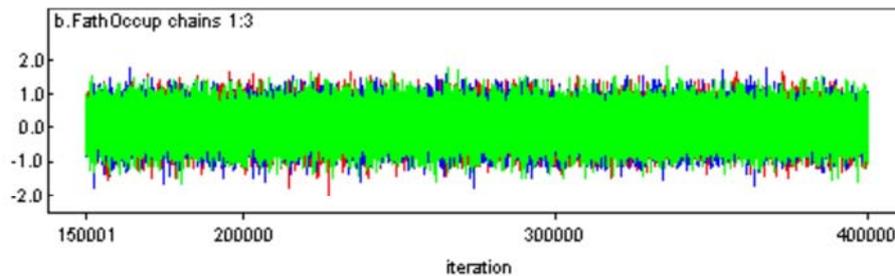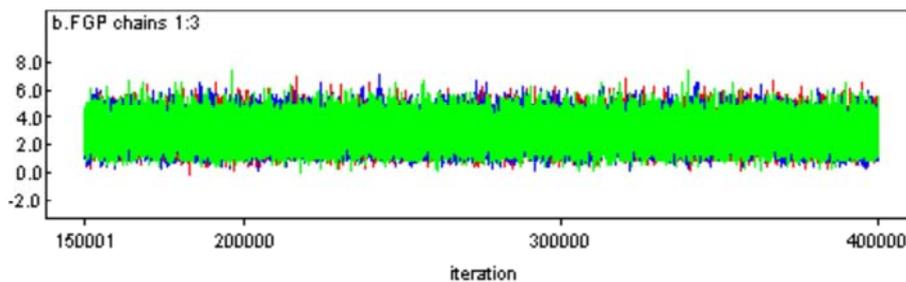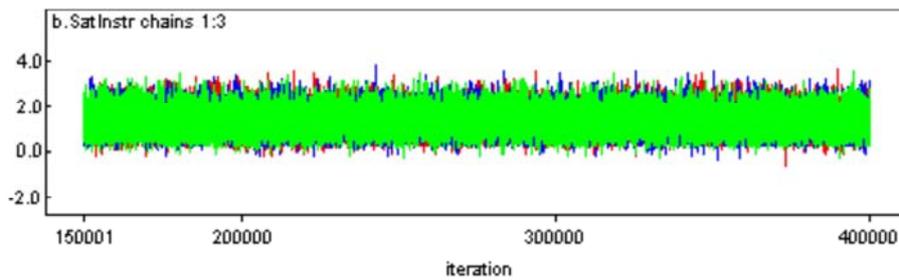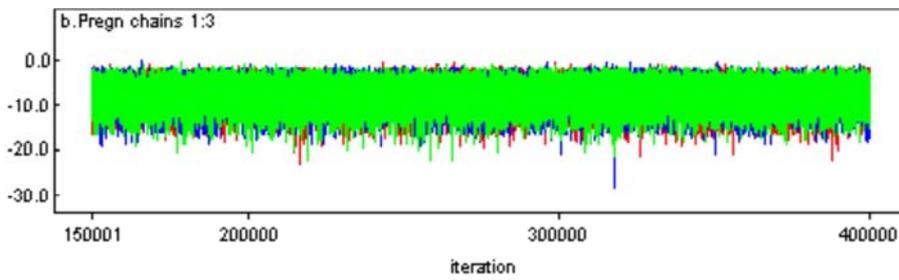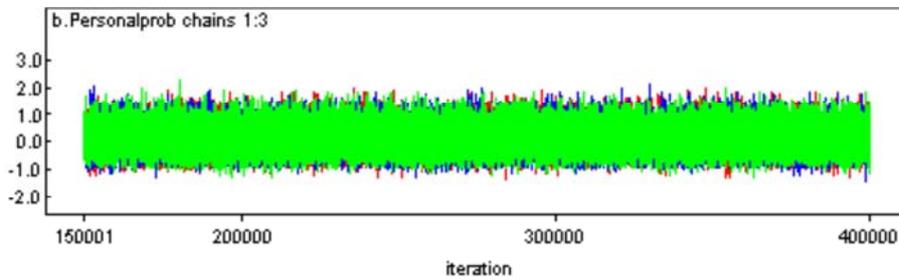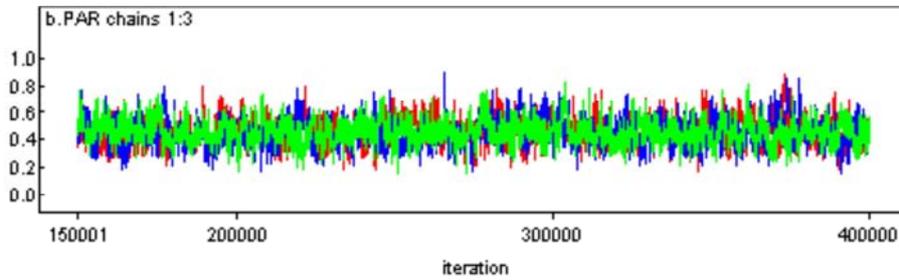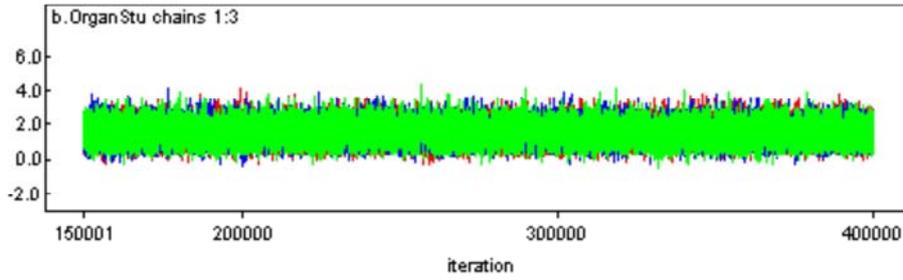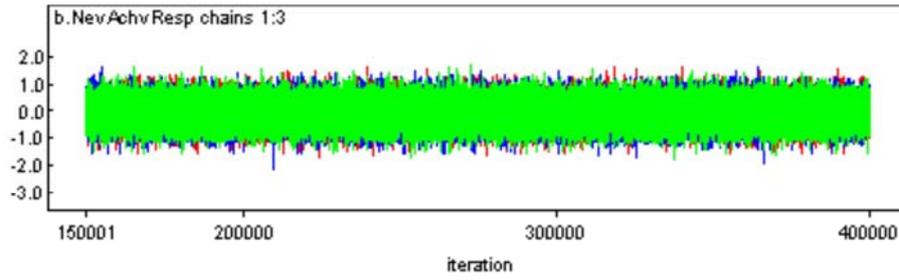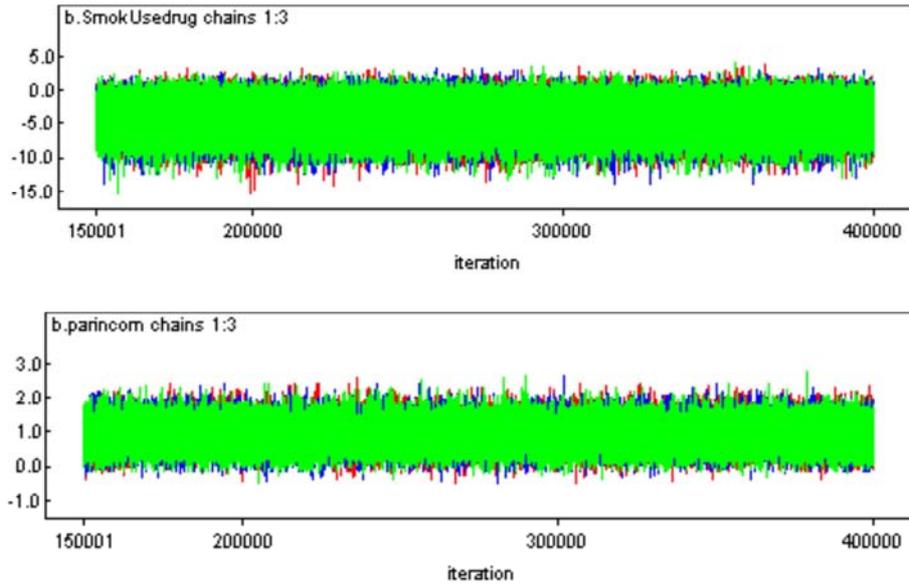
Figure 5. *Convergence of Time Series Plots for the Coefficients in the Retention Analysis Data.*

b. Autocorrelation Plot for retention data:

The plots showed that the three independent chains were mixed or overlapped to each other and died out for higher lags and hence this is an evidence of convergence (See Figure 6).
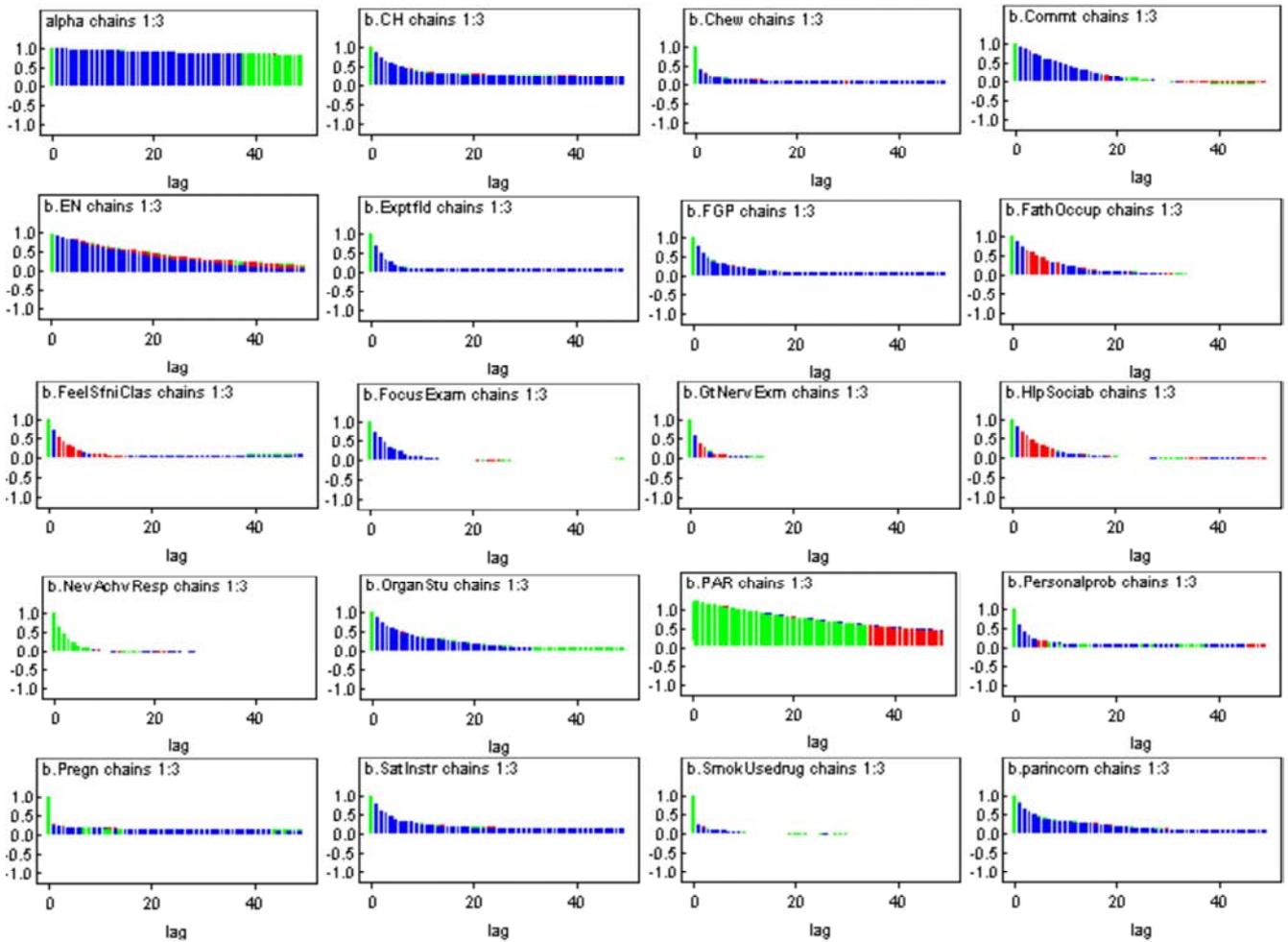

Figure 6. *Convergence based on autocorrelation plots for the Coefficients in the Retention Analysis Data.*

c. Gelman-Rubin statistic (GR) for retention data:

Evidence for convergence comes from the red line being close to 1 on the y-axis and from the blue and green lines being stableacross the width of the plot [26]. Hence, evidence for convergence has been reached (See in Figure 7).
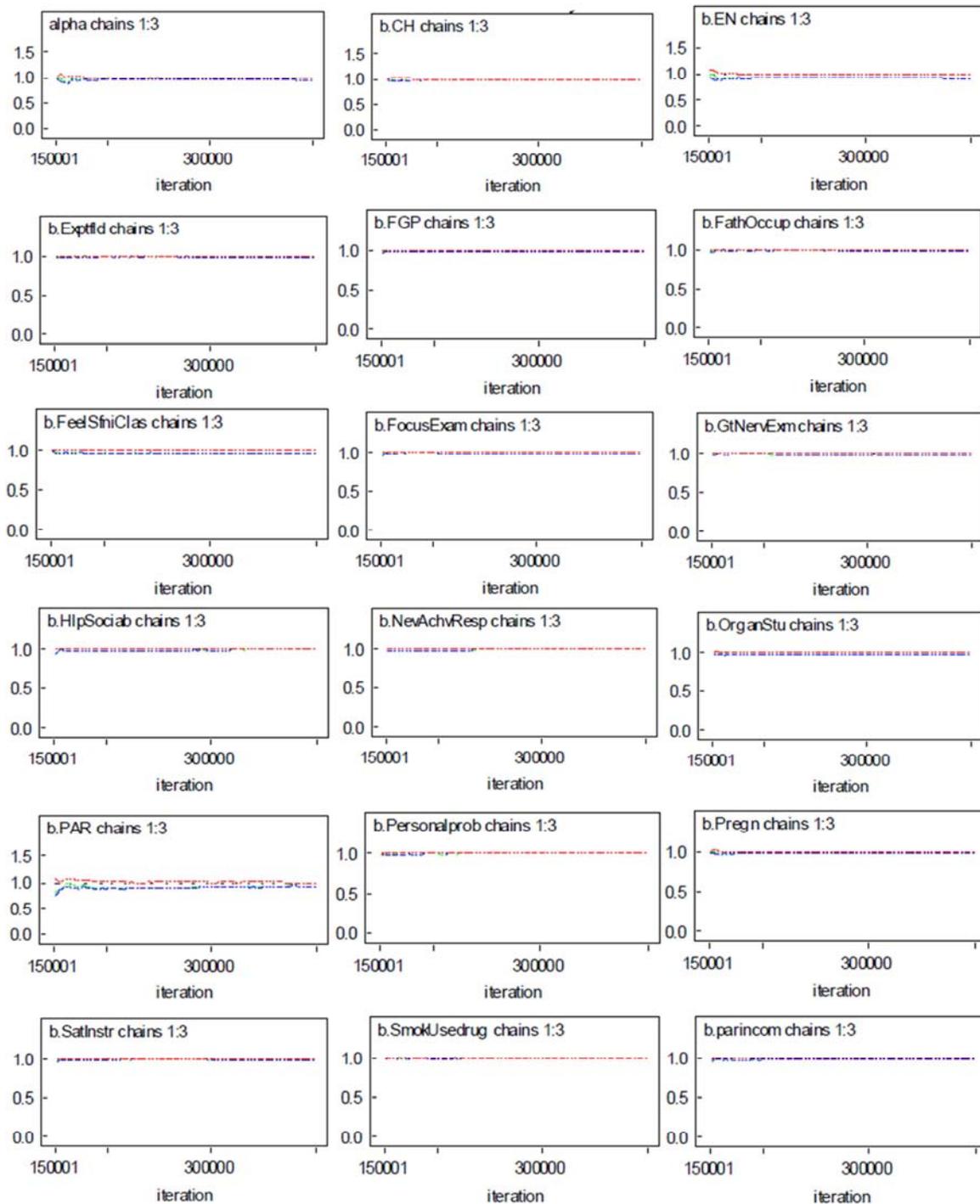

Figure 7. Convergence using Gelman-Rubin Statistic for the parameters in the Retention Analysis Data.

d. Density Plot for retention data:

Since the plots for the most predictor variables indicated that the coefficient has normal distribution, the simulated parameter value indicated convergence (See Figure 8)
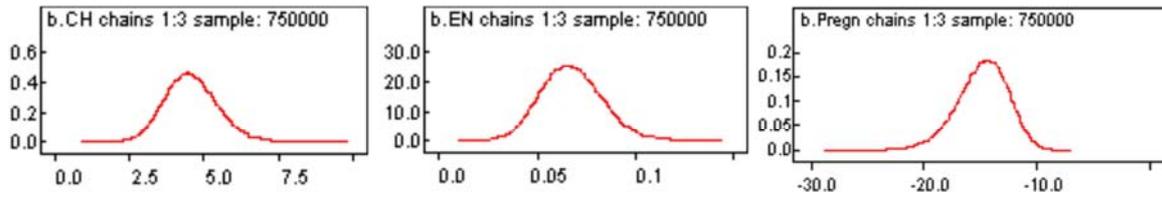
Figure 8. *Convergence for Density Plot of the Parameters of Choice of Field (b. CH), Entrance Exam Score (b. EN), and pregnancy (b. Pregn) in the Retention Analysis Data.*

ii. Assessing Accuracy of Bayesian Logistic Regression Model for Retention Data.

Once convergence has been achieved, we need further simulation for a further number of iterations to obtain samples that can be used for posterior inference. Simulation should be run until the Monte Carlo error for each parameter of interest is less than about 5% of the sample standard deviation.

Table 12 contains the estimated coefficients, Mean ($\hat{\beta}$), the standard deviation sd, Monte Carlo (MC) errors, 5% of the standard deviation (0.05*sd) and 95.5% Confidence interval.

The MCE for each significant predictor variable is less than 5% of its posterior standard deviation and hence convergence and accuracy of posterior estimates are attained and the model is appropriate to estimate posterior statistics. The confidence intervals for the constant (alpha) and for the predictor variables; choice of field (b. CH), habit of chewing chat (b. Chew), entrance exam score (b. EN), first year CGPA (b. FGP), feel safe to study at night in classrooms (b. FeelSfniClas), organizing study and leisure time (b. OrganStu), preparatory average result (b. PAR), pregnancy (b. Pregn), satisfaction withinstructors (b. SatInstr), habit of

smoking cigarette and using drugs (b. SmokUsedrug) and parent income (b. parincom) does not include zero indicating they are significant. Therefore, these predictor variables mainly affect female students' retention. The Mean ($\hat{\beta}$), values for habit of chewing chat (b. Chew), feel safe to study at night in classrooms (b. FeelSfniClas), pregnancy (b. Pregn) and smoking cigarette and using drugs (b. SmokUsedrug) are negative indicating negative relationship between the dependent variable retention and these predicator variables. While those values for the other significant predictors were positive indicating positive relationship. For example, in the case of choice of field (b. CH), since 'No' is coded as 0 and 'Yes' is coded as 1, the positive value indicated that female students who assigned to the field they chose were more likely to retain than those didn't assigned to their field choice by their interest. In the case of pregnancy (b. pregn), since 'No' is coded as 0 and 'Yes' is coded as 1, the negative value indicated that female students who ever got pregnancy were less likely to retain than those did not get pregnancy (Results are displayed in Table 12).

Table 12. *Summary Statistics of the Posterior Distribution of the Model Parameters Obtained from the Sampled Values (HU, 2008/09 Entries).*

Variables	Mean ($\hat{\beta}$)	sd	MC error	0.05*sd	95.5% CI	
					Lower	Upper
alpha	-23.47	5.003	0.0940	0.250	-34.07	-14.43*
b. CH	4.566	0.887	0.0096	0.044	2.945	6.423*
b. Chew	-3.507	1.142	0.0077	0.057	-5.826	-1.339*
b. Commt	-0.506	0.601	0.0030	0.030	-1.723	0.6357
b. EN	0.068	0.016	0.0003	0.001	0.039	0.1012*
b. Exptfld	0.386	0.365	0.0023	0.018	-0.324	1.112
b. FGP	3.004	0.749	0.0052	0.038	1.609	4.548*
b. FathOccup	0.084	0.375	0.0017	0.019	-0.665	0.8086
b. FeelSfniClas	-0.89	0.409	0.0025	0.021	-1.724	-0.1182*
b. FocusExam	0.148	0.398	0.0018	0.020	-0.630	0.9319
b. GtNervExm	-0.398	0.356	0.0010	0.018	-1.11	0.2911
b. HlpSociab	0.261	0.43	0.0015	0.022	-0.585	1.105
b. NevAchvResp	-0.039	0.363	0.0010	0.018	-0.756	0.671
b. OrganStu	1.59	0.481	0.0042	0.024	0.691	2.58*
b. PAR	0.458	0.092	0.0020	0.005	0.292	0.6544*
b. Personalprob	0.243	0.377	0.0023	0.019	-0.482	0.9978
b. Pregn	-7.705	2.212	0.0181	0.111	-12.47	-3.834*
b. SatInstr	1.42	0.427	0.0037	0.021	0.627	2.303*
b. SmokUsedrug	-3.824	1.879	0.0041	0.094	-7.77	-0.375*
b. parincom	0.937	0.334	0.0026	0.017	0.306	1.617*

3.2.4. Interpretation and Discussion for the Retention Analysis Data Results

This study has provided some information about the factors those determine female students' retention at HU. According to the results, among the 243 female students who were 2008/09 entry with mean age at first year equals 18.5761, 75.7%

retained but the rest 24.3% did not. Of which, 47.7% were from the Natural Science Faculty, 32.9% were from the Social Science Faculty and 19.3% were from the Agriculture Faculty. In the case of choice of field, 'Yes' stands for the 1st or 2nd choices but 'No' stands for the student's 3rd or above choices. Hence 79.4% were assigned to their first or second

choices and the rest 20.6% were assigned to their 3rdor above choices. About 77% of them were committed to their academic activities. About 71.6% of female students were able to organize their study and leisure time easily. About 40.7% of them were being nervous and forgot facts they know on exam and 77.8% of them were having a CGPA of greater than or equal to 2.0 but the rest 22.2% of them were having a CGPA of less than 2.0. About 53.5% of female students were satisfied with their instructors.

From the bivariate and Univariate results, choice of field, first year CGPA, pregnancy, parent income, habit of chewing chat, habit of smoking cigarette and using drugs, focus on exam, commitment to academic activities, organize study and leisure time, satisfaction with instructors, never achieve academic respects, father occupation, being nervous and forget facts on exam, help of sociability in campus to reach educational goals, feel safe to study at night in classrooms, expectation about field and weak because of personal problems were found significant predictors of female students' retention.

From the Bayesian logistic regression analysis, choice of field (b. CH), habit of chewing chat (b. Chew), entrance exam score (b. EN), first year CGPA (b. FGP), feel safe to study at night in classrooms (b. FeelSfniClas), organizing study and leisure time (b. OrganStu), preparatory average result (b. PAR), pregnancy (b. Pregn), satisfaction with instructors (b. SatInstr), habit of smoking cigarette and using drugs (b. SmokUsedrug) and parent income (b. parincom) were found significant predictors of female students' retention. The Mean ($\hat{\beta}$), values for habit of chewing chat (b. Chew), feel safe to study at night in classrooms (b. FeelSfniClas), pregnancy (b. Pregn) and smoking cigarette and using drugs (b. SmokUsedrug) were negative indicated negative relationship between the dependent variable retention and these predicator variables. However, those values for the other significant predictors were positive indicated positive relationship.

4. Conclusions

The objective of this study was to identify the factors those affect female students' graduation and retention.

Graduation of female students depends significantly upon several factors. The bivariate and univariate analyses indicated that preparatory average result, choice of field, entrance exam score, first year CGPA and age significantly affect female students' graduation. The Bayesian logistic regression analyses showed that preparatory average result, choice of field, entrance exam score and first year cumulative GPA mainly affected female students' graduation. The graduation rate for female students who assigned to the field they chose by their interest was higher than those assigned to the field they did not chose. Female students whose first year CGPA less than 2.0 were having lower rate of graduation than those having first year CGPA greater than 2.0. Graduation rate for female students having higher preparatory average result was higher than that for female students having lower preparatory average result.

Retention of female students also depends significantly upon several factors. Here, the Bayesian logistic regression analyses showed that preparatory average result, choice of field, first year CGPA, pregnancy, satisfaction with instructors, habit of chewing chat, parent income and organize study and leisure time mainly affected female students' retention. Entrance exam score, feel safe to study at night in classrooms and habit of smoking cigarette and using drugs were also significant. The retention rate for female students who assigned to their 1stor 2ndchoice was higher than those assigned to their 3rdor above choice. Female students whose first year CGPA less than 2.0 were having lower rate of retention than those having first year CGPA greater than 2.0. Retention rate for female students having higher preparatory average result was higher than that for those having lower preparatory average result. Female students having parents' income less than 500 were less likely to retain than those having parents' income greater than 1500. The retention rate for female students who were satisfied with their instructors was higher than those not satisfied. Female students who can organize their study and leisure time easily were more likely to retain than those cannot. Female students who have a habit of chewing Khat and also who have a habit of smoking cigarette and using drugs were less likely to retain than those have no. The retention rate for female students who ever got pregnancy was lower than that for those who did not. The retention rate for female students who feel safe to study at night in classrooms was lower than that for those who did not.

We recommended that designing an appropriate method of enrolment and allocation of students to different faculties and departments may help female students' graduation and retention. Campus and Department officers in collaboration with the students themselves and academic staffs are expected to work hard to bring change in behaviour, academics and social aspects of female students. Special attention must be given to the task of female students advising in order to help integrate them both socially and academically into the university environment. We also recommend that the teaching method at secondary and preparatory schools should be designed to challenge students and motivate them to prepare adequately for Higher Education Institutions.

Acknowledgements

References

[1] Rose, P. (2003). Can Gender Equality in Education Be Attained?: Evidence from Ethiopia Background Paper for 2003 UNESCO Global Monitoring Report. Centre for International Education, University of Sussex.

[2]   Asresash Demise, Ruth Shinebaum, KassahunMelesse. (2002). The Problems of Female Students at Jimma University, Ethiopia, with Some Suggested Solutions. Ethiopian Journal of Health, 16 (3): pp 257-266.

[3]   Carmel, A., Gold S. (2007). The effects of course delivery modality on student satisfaction and retention and GPA in on-site vs. Hybrid courses. Journal of Distance Education, 8 (2), 1302–6488.

[4]   Foster, K. (2003). Libraries and Student Retention: Some Thoughts about the Issues and an Approach to Evaluation. Journal of SCONUL Newsletter: 28, pp. 12-16.

[5]   Stephenson, C., Derbenwick Miller, A., Alvarado, C., Barker, L., Barr, V., Camp, T., Frieze, C., Lewis, C., Cannon Mindell, E., Limbird, L., Richardson, D., Sahami, M., Villa, E., Walker, H., and Zweben, S. (2018). Retention in Computer Science Undergraduate Programs in the U. S.: Data Challenges and Promising Interventions. New York, NY. ACM.

[6]   Ferreira, MM., (2009). "Trends in women's representation in science and engineering", Journal of Women and Minorities in Science and Engineering, vol/issue: 15 (3), pp. 191-203.

[7]   Mastekaasa, A., Smeby, J., (2008) "Educational choice and persistence in male and female dominated fields", Higher Education, vol/issue: 55 (6), pp. 189-202.

[8]   Military Leadership Diversity Commission. (2010). Differences in Promotion and Retention Rates by Race/Ethnicity and Gender: Considerations When Interpreting Overall Continuation Rates [Issue Paper #30]. Arlington, VA: Military Leadership Diversity Commission.

[9]   Mississippi University for Women website. (2019). MUW Graduation Rate & Retention Rates. Accessed from https://www.collegefactual.com/colleges/mississippi-university-for-women/academic-life/graduation-and-retention/ on May 15, 2019.

[10]   Zhang, G., Anderson, T. J., Ohland, M. W., Carter R. and Thorndyke, B. (2004). Identifying Factors Influencing Engineering Student Graduation: A Longitudinal and Cross-Institutional Study. Journal of Engineering Education, 93 (4), pp. 313-320.

[11]   Toolkit for Women. (2005). United Nations 3rd World Conference Nairobi. Accessed from www. Earthsummit2002.org/toolkits/women/un-doku/un-conf/nairobi.htm.

[12]   KassaShurke. (2008). Factors Affecting Females' Participation in Education: The Case of TochaWoreda in SNNPR. Addis Ababa University. Accessed from http://hdl.handle.net/123456789/876.

[13]   Greenberger, M. D., Samuels, J., Chaudhry, N. K., Graves, F. G., Kaufmann, L., Keeley, T., Null, T. W., Seemeyer, L., Gupta, K., Kuznick, A., Petronko, K., Robinson, K., White, E., and Wong, K., with LeMair, L. M. (2007). When Girls Don't Graduate We All Fail. National Women's Law Center. Washington, DC 20036, www.nwlc.org.

[14]   Seyoum, Tefera. (1991). The Participation of Girls in Higher Education in Ethiopia: In the Proceeedings of the First University Seminar on Gender Issues in Ethiopia held in December 24-26. pp. 99-108. Institute of Ethiopian Studies, Addis Ababa University.

[15]   DemewezAdmassu, MehadiAbdo and TesfayeSemela. (2005). Impact of Varying Entry Behaviour on Students" Academic and Psychological Outcomes in Higher Education: The Case of PPC and FPC Students at Debub University. The Ethiopian Journal of Higher Education: 2 (2), pp 49-86.

[16]   TesfayeSemela. (2005). Higher Education Expansion and the Gender Question in Ethiopia: A Case Study of Women in a Public University. Proceedings of the Conference on the Future Direction of Higher Education in Ethiopia.

[17]   Mezick, E. M. (2007). Return on Investment: Libraries and Student Retention. Journal of Academic Librarianship: 33 (5), pp. 561-566.

[18]   Bean, J. and Eaton, S. B. (2001). The Psychology Underlying Successful Retention Practices. Journal of College Student Retention: Research, Theory and Practice. 3 (1), pp. 73 – 89.

[19]   Felder, R. M., Felder, G. N., Mauney, M., Hamrin, C. E., Jr., & Dietz, E. J. (1995). A Longitudinal Study of Engineering Student Performance and Retention. III. Gender Differences in Student Performance and Attitudes. Journal of Engineering Education: 84 (2), pp. 151–174.

[20]   Gupta, K. (2008). When Girls Don't Graduate, We all Fail: Helping Girls Stay in School: Spotlight on Pregnant and Parenting Teens: NAPE/Women Work! Conference, Taryn Wilgus Null, National Women's Law Center, Abby Kahn, Healthy Teen Network.

[21]   Central Statistical Agency (CSA). (2007). Demographics of Ethiopia.

[22]   Cochran, W. G. (1997). Sampling Techniques. (3rd Ed.). John Wiley and Sons, Inc, New York.

[23]   Al-Subaihi, A. A. (2003). Sample Size Determination: Influencing Factors and Calculation Strategies for Survey Research. Journal of Saudi Med J: 24 (4), pp. 323 – 330. Accessed from www.smj.org.sa, on July 5, 2010.

[24]   Stamps, D. K. (2006). Markov Chain Monte Carlo Methods for Regression Splines with a Penalized Acceptance Ration. Dissertation. 600. University of Missouri, St. Louis.

[25]   Geyer, C. J. (1992). On the convergence of Monte Carlo maximum likelihood calculations. Technical report 571, School of Statistics, University of Minnesota.

[26]   Merkle, E., Sheu, C., and Trisha, V. Z. (2005). Simulation – Based Bayesian Inference Using WinBUGS.