Modelling of the Advanced Level National Examination Average Pass Rate in Zimbabwe using Bayesian Hierarchical Log-logistic and Normal Mixture Approach

Barbara Ngwarati 1*, Nur Iriawan 1, and Heri Kuswanto 1
1Institut Teknologi Sepuluh Nopember (ITS), Surabaya 60111, Indonesia

*Corresponding author: barbarangwarati@gmail.com

Abstract. The national examination as one of the standard evaluation systems of education in Zimbabwe is used for the educational developments that seek to improve the quality of education in the educational sectors. This research aims to find the best model and its factors affecting the average pass rate of the Advanced Level (A-Level) national examination in Zimbabwe. Modelling was conducted using a two-level hierarchical model with factors influencing the national examination at district in the first level and those influencing the national examination provincial level in the second level. The Bayesian approaches namely hierarchical log-logistic and normal mixture were used in the modelling. The estimation of these Bayesian approaches posterior parameters was done using Markov Chain Monte Carlo (MCMC) and the Deviance Information Criterion (DIC) value was used to select the best model. The hierarchical normal mixture was found to be the best model to explain the variability of the average pass rate percentage of the A–Level national examination and all the micro and macro variables in this study significantly influenced the A-Level national examination in Zimbabwe.

1. Introduction

National Examinations play a major role in the educational system of Zimbabwe. National examinations in Zimbabwe are administered by a board known as Zimbabwe School Examinations Council (ZIMSEC). This board is responsible for assessment in primary and secondary education in Zimbabwe. ZIMSEC administers the national examinations at the end of 7 years primary course and at the end of the secondary course which is divided into 4 years in junior high school and senior 2 years in high school. The role of national examinations in secondary education is important in assessment and evaluation on what is taught in schools, facilitation of entry to schools, selection of students during their careers and students performance evaluation which is very important for their further education and even life choices [1].

ZIMSEC's other role is to assess the national exam pass rate at primary and secondary level so that the areas which need improvement can be addressed hence engaging in this study. National examination data often has a hierarchical structure which is represented by the sub-models. These sub-models illustrate the connections between variables in the first level and explain the effect of connections with
variables at higher levels known as the micro model and macro model [2] and [3]. The micro model is the model at the lowest hierarchical level structure while the macro model is at the highest. In this study, the national examination data consists of two levels namely the micro level which is the model at the district level whilst the macro level is the model found at the provincial level. The 2017 Advanced Level (A-Level) average percentage pass rate in Zimbabwe indicated a multi-modal data pattern and moreover consists of a two-level hierarchical model. This study seeks to investigate the work on Bayesian hierarchical log-logistic and normal mixture model combined with the Markov Chain Monte Carlo (MCMC) approaches to model the A-level average percentage pass rate in Zimbabwe.

2. Theoretical Review

2.1 Identification of the Mixture in the Model
In the mixture model, the data has multimodal properties which consist of subpopulations and each subpopulation has different proportions. One of the methods used to detect the existence of the mixture in the data distribution pattern is the histogram method [3]. Needless to say, this method determines the number of mixture components by counting the number of peaks on the histogram. Two mixture components in this study were determined using the histogram method.

2.2 Identification of the Distribution of Data
The distribution of data is identified through carrying out the goodness of fit test. In this study, the Kolmogorov-Smirnov test was carried out to test the matching between the observation data to the selected distribution. The A-level national examination pass rate percentage data in the districts of Zimbabwe followed both the three-parameter log-logistic and normal distribution with p values of 0.7436 and 0.43852 respectively.

2.3 Log-logistic and Normal Distribution
The log-logistic distribution of three-parameters is a probability distribution also known as the generalized log-logistic or shifted log-logistic distribution. Given a random variable \( t \) which has a two-parameters log-logistic distribution with shape, scale and location parameters as \( \mu > 0, \sigma^2 > 0 \) and \( \theta = 0 \) respectively, denoted by \( t \sim LL (\mu, \sigma^2, \theta) \) then \( Y = \log (t) \) has a logistic distribution. This means that \( \exp(Y) \sim LL (\mu = e^\theta, \sigma^2 = \frac{1}{s}, \theta) \) leads to a three-parameters log-logistic \( \exp(Y) + \theta \sim shifted (\mu = e^\theta, \sigma^2 = \frac{1}{s}, \theta) \) where \( s \) is the scale parameter. The domain of the log-logistic three-parameters distribution is given by \( \theta \leq t < \infty \) and its PDF is given by [5]:

\[
f(t, \mu, \sigma^2, \theta) = \frac{2}{\sqrt{2\pi}} \exp\left(\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) (1 + \left(\frac{t-\mu}{\sigma}\right)^2)^{-\frac{3}{2}}
\]

The standard normal distribution is given by the variance, \( \sigma^2 \) and mean, \( \mu \) and in short form, it is written as \( N (\mu, \sigma^2) \). The domain of a normal distribution is \( -\infty < t < +\infty \) and the PDF of the normal distribution is given as:

\[
f(t, \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{t-\mu}{\sigma}\right)^2}
\]

2.4 Hierarchical Linear Model
Hierarchical linear models are designed to analyze data with a variety of complex sources [6]. The hierarchical model consists of sub-models which are in levels with the micro model at the first level and the macro model in the second level. These sub-models have a connection with each other. For a two levels hierarchical data structure in this study, it consists of \( m \) groups (i.e the number of provinces) and in each group, there consists of \( N_j \) (i.e. the number of districts in \( j \)th provinces) where \( j = 1, 2, \ldots, m \).
The number of the response variable for the $j^{th}$ group for each $N_j$ is given as $Y_{ij}, Y_{2ij}, \ldots, Y_{N_{ij}}$. On the other hand, $X_{1ij}, X_{2ij}, \ldots, X_{ pij}$ are predictor variables at the first level for the $j^{th}$ group and $G_i, G_2, \ldots, G_q$ are predictor variables at the second level. [7] and [8] in [9] used equation (3) to describe the first level of a two-level hierarchical model where $i = 1, 2, \ldots, N_j$;

$$Y_{ij} = \beta_0 + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \cdots + \beta_{pj}X_{pj} + e_{ij}$$

(3)

which can be represented in matrix form as shown in equation (4):

$$y_j = X_j \beta_j + e_j$$

(4)

The regression models $m$ in the micro model will produce many regression parameters of value $\beta_j$ that varies between groups where $r = 0, 1, \ldots, p$ with $r$ as the number of regression parameters and $p$ as the number of micro predictor variables. The equation of the macro model is given as follows:

$$\beta_{rj} = \gamma_{0r} + \gamma_{1r}G_{1j} + \gamma_{2r}G_{2j} + \cdots + \gamma_{lr}G_{lj} + u_{rj}$$

(5)

or can be expressed as:

$$\beta_r = G_r \gamma_r + u_r$$

(6)

[8] explains that the combination of the micro and macro equations results in a combined equation (7) as follows:

$$y_{ij} = y_{00} + \sum_{r=1}^{p} y_{or} x_{r(i)} + \sum_{q=1}^{s} y_{q0} G_{qi} + \sum_{q=1}^{s} \sum_{r=1}^{p} y_{rq} G_{qj} x_{r(i)} + u_{0j} + \sum_{r=1}^{p} u_{rj} x_{r(i)} + e_{ij}$$

(7)

which can be written in vector form:

$$y_j = X_j G_j \gamma + X_j u_j + e_j$$

(8)

with, $X_j G_j \gamma$ and $X_j u_j$ representing the fixed tribes and random tribes respectively in the hierarchical model, $E(y_j) = X_j G_j \gamma$ and $Var(y_j) = X_j^T X_j + \sigma^2 I_{nj}$.

2.5 Log-logistic and Normal Mixture Model

Given $S$ sub-populations of each is log-logistically or normally distributed, the density function of the $j^{th}$ mixture model of the log-logistic and normal distribution are given as follows respectively [6]:

$$p(y | \pi, \alpha, \beta, \gamma) = \sum_{s=1}^{S} \pi_s g_s(x | \alpha, \beta)$$

and

$$p(y | \pi, \mu, \sigma^2) = \sum_{s=1}^{S} \pi_s g_s(x | \mu, \sigma^2)$$

(9)

where $p(y | \pi, \alpha, \beta, \gamma)$ and $p(y | \pi, \mu, \sigma^2)$ are density function of the log-logistic and normal distribution mixture model respectively, $g_s = (y | \alpha_s, \beta_s, \gamma_s)$ and $g_s = (y | \mu_s, \sigma^2_s)$ are the log-logistic and normal distribution $S^h$ density function respectively, where $s = 1, 2, \ldots, S$ is the mixture components, $(\alpha_s, \beta_s, \gamma_s)$ and $(\mu_s, \sigma^2_s)$ are the vector parameters for log-logistic and normal distribution respectively, $\pi_s$ is the proportion of each component in the mixture model with elements $\pi = \{\pi_1, \pi_2, \ldots, \pi_s\}$, $\pi_1 + \pi_2 + \cdots + \pi_s = 1$ and $\pi_s$ lie between 0 and 1 exclusively.

2.6 Bayesian Theory to the Distribution of the Mixture

Nowadays Bayesian methods are used for complex model analysis in statistics [10]. The methods view the unknown parameters as random variables which are characterized by the parameters of the prior distribution [11]. The estimation of the posterior parameters involves the combination between the prior distribution and the observational data (likelihood function) [12]. The Bayes' theorem states that if there are $\theta$ parameter obtained by observation data $y$ hence the posterior $\theta$ is expressed as follows:

$$p(\theta | y) \propto p(y | \theta) \ p(\theta)$$

(10)

wherein this case the total probability as a normalized constant amount to one [13]. The informative priors used in hierarchical log-logistic and normal mixture modelling was pseudo and conjugate.
The likelihood of a mixture model, \( l_{\text{mix}} \) for log-logistic and normal distribution respectively are shown in equation (11) with \( n_1 + n_2 + \ldots + n_s = n \) where \( S \) is the number of mixture components in \( n \) observed data.

\[
l_{\text{mix}} = \prod_{i_1=1}^{n_1} p_1(y_{i_1}|\pi, \alpha, \beta, \gamma) + \prod_{i_2=1}^{n_2} p_2(y_{i_2}|\pi, \alpha, \beta, \gamma) + \ldots + \prod_{i_s=1}^{n_s} p_s(y_{i_s}|\pi, \alpha, \beta, \gamma)
\]

and

\[
l_{\text{mix}} = \prod_{i_1=1}^{n_1} p_1(y_{i_1}|\pi, \mu, \sigma^2) + \prod_{i_2=1}^{n_2} p_2(y_{i_2}|\pi, \mu, \sigma^2) + \ldots + \prod_{i_s=1}^{n_s} p_s(y_{i_s}|\pi, \mu, \sigma^2)
\]

2.7 Combined Mixture Model.

The micro and macro model formed a two-level hierarchical mixture model. In this study, the micro model is a single equation model at the first level formed with the log-logistic or normal regression mixture models. On the other hand, the macro model is another single equation model at the second level. The combined model in either equation (7) and (8) is combined with the log-logistic and normal mixture components from (9) to obtain (12) in each instance to produce the combined mixture model, \( \mu_B \).

\[
\mu_B = \sum_{s=1}^{S} \pi_s (X_j G_j \gamma + X_j \mu_f + e_f)
\]

2.8 Markov Chain Monte Carlo Algorithm (MCMC)

Posterior distributions in Bayesian analysis require a method of the numerical approach which is run iteratively until convergence known as the MCMC [14,15]. A parameter significance test was used to identify significant parameters in the model through the concept of the credible interval. The null hypothesis states that the parameter is significant while the alternative hypothesis states that the parameter is not significant. Hence the test checks for the inclusion of zero within the credible intervals and if zero is included then the null hypothesis is rejected hence concluding that the parameter is not significant [16].

2.9 Deviance Information Criterion (DIC)

DIC is used to measure the depth of the adequacy of a model or make comparisons in between or among models [17,18]. The lowest DIC is preferable in selecting the best model.

3. Results and Discussions

3.1. Data Source and Research Variables

The data used in this study is secondary data and was obtained from the Zimbabwe Ministry of Education, Sports and Culture. The data is on the average percentage pass rate in all the districts of Zimbabwe for the November 2017 national examinations. The sample at the first level was a total of 72 districts and a sample of 10 provinces in the second level. The response variable was the A-Level average percentage pass rate (Y), the micro variables were factors affecting the pass rate at the district level (X) and the macro variables were factors affecting the pass rate at the provincial level (G). Table 1 shows the number of variables of this study at both levels.

There were six micro variables and eight macro variables, and this study involves identification of the variables that have a significant effect on the average pass rate in Zimbabwe and modelling of the A-Level average pass rate percentage using the Bayesian hierarchical log-logistic and normal mixture by employing the WinBUGS software. Specifically, for this study, the provinces in Zimbabwe were labeled from 1-10 in the order which includes Bulawayo, Harare, Manicaland, Mashonaland Central, Mashonaland East, Mashonaland West, Masvingo, Matabeleland North, Matabeleland South, and Midlands provinces respectively.
Table 1. Number of variables in the two-level hierarchical model

| First Level Variables | Second Level Variables |
|-----------------------|------------------------|
| X₁: Number of public libraries | G₁: Percentage of repeaters |
| X₂: Number of orphaned and vulnerable learners | G₂: Percentage of schools without electricity |
| X₃: Percentage of qualified teachers | G₃: Average learners classroom ratio |
| X₄: Teacher to learner ratio | G₄: Percentage of schools without a water source |
| X₅: Number of pupils on financial aid | G₅: Average learners per computer ratio |
| X₆: Classroom to teacher ratio | G₆: Number of pupils without seating facilities |
|                       | G₇: Number of pupils without writing facilities |
|                       | G₈: School participation rate |

3.2 Identification of the distribution of the mixture

The A-level pass rate data distribution was done using a histogram and the data had a multimodal pattern. Table 2 shows that the A-Level average pass rate data has two modes, therefore, assuming that the data comes from multi-modal distributions.

Table 2. A-level Average Pass Rate Descriptive Statistics

| Variable                                | Mean | Minimum | Maximum | Mode    |
|-----------------------------------------|------|---------|---------|---------|
| A-level Average Pass Rate (%)           | 85.5 | 64      | 93.4    | 82.9, 91.3 |

The average pass rate of the national examination from 72 districts in Zimbabwe for the year 2017 was 85.5% with the lowest value of 64% and the highest value of 93.4%. The data had no missing values and had two modes namely 82.9% and 91.3%, leading to a conclusion that the average national examination data was multimodal. The data had 2 mixture components, the first component had 5 samples and the second component had 67 samples. The average percentage pass rate in mixture 1 and in mixture 2 were 75.4% and 86.3% respectively. This meant that the pass rate was higher in the second mixture as compared to the first one.

3.3 Implementation of Log-logistic and Normal Hierarchical Mixture Model on the A-Level Average Pass Rate

The modelling and identification of the factors affecting the national examination pass rate were done by using the following four models and the parameter significance test was used.

3.3.1 Log-logistic Hierarchical Mixture Model

Modelling was done with 6 variables at level 1 and 8 variables at level 2. Two mixture components were considered and 70 parameters in level 1 and 126 parameters in level 2 were estimated. The iteration used 136000 times with 100 thin so that 136000 samples were used to estimate the parameters.

Table 3 shows two significant parameters in mixture 1 and some significant parameters in mixture 2 from a total of nine micro model significant parameters of the log-logistic mixture model. For example, in the first row in Table 3, the mixture 1 parameter \( \beta_{1.2} \) indicates the first regression parameter in the second province and its mean, standard deviation, MC error, and median are -856.9; 479.6; 24.94 and -920.8 respectively. The mean indicates the significance of the parameters as it lies within the credible
interval (i.e. 2.5% to 97.5%) that stretches from -1608 to -25.84 and this pattern is the same for each significant parameter in Table 3. Moreover, the absence of the zero value within each credible interval proves the positivity of the parameter significant test. All significant parameters were obtained in mixture 1 and mixture 2, for example, the parameter $\beta_{1,2}$ is associated with the number of public libraries which belongs to Harare province and the parameter $\beta_{5,2}$ is associated with pupils who depend on financial aid also in the Harare province.

Table 4 shows some of the significant macro parameters in mixture 2. However, there were thirteen significant parameters in total and these were found in mixture 2 only since there were no significant parameters in mixture 1. The variables which were significant in mixture 2 include the percentage of schools without electricity, percentage of schools without a water source, average learners' classroom ratio, average learners per computer ratio, number of pupils without seating facilities, number of pupils without writing facilities and lastly school participation rate except for the number of repeaters variable.

Table 3. Significant parameters of the log-logistic mixture micro model in mixture 1 and mixture 2

| Mixture | Parameter | Mean  | SD    | MC Error | 2.5%   | Median  | 97.5%   |
|---------|-----------|-------|-------|----------|--------|---------|---------|
| 1       | $\beta_{1,2}$ | -856.9 | 479.6 | 24.94    | -1608  | -920.8  | -25.84  |
| 1       | $\beta_{5,2}$ | -856.9 | 479.6 | 24.94    | -1608  | -920.7  | -25.8   |
| 2       | $\beta_{0,2}$ | 4.34   | 0.97  | 0.003    | 2.41   | 4.34    | 6.25    |
| 2       | $\beta_{0,3}$ | 4.50   | 0.83  | 0.003    | 2.86   | 4.50    | 6.17    |
| 2       | $\beta_{0,4}$ | 4.50   | 1.02  | 0.003    | 2.45   | 4.49    | 6.55    |
| ...     | ...       | ...   | ...   | ...      | ...    | ...     | ...     |
| 2       | $\beta_{0,9}$ | 4.58   | 1.27  | 0.004    | 2.04   | 4.58    | 7.10    |
| 2       | $\beta_{0,10}$ | 4.43   | 0.53  | 0.001    | 3.37   | 4.43    | 5.50    |

Table 4. Significant parameters of the log-logistic mixture macro model in mixture 2

| Parameter | Mean  | SD    | MC Error | 2.5%   | Median  | 97.5%   |
|-----------|-------|-------|----------|--------|---------|---------|
| $\gamma_{0,0}$ | 90.25 | 1.308 | 0.004    | 87.68  | 90.25   | 92.81   |
| $\gamma_{2,0}$ | 31.25 | 6.148 | 0.017    | 19.15  | 31.26   | 43.28   |
| $\gamma_{3,0}$ | 10.96 | 4.014 | 0.011    | 3.09   | 10.97   | 18.87   |
| $\gamma_{4,0}$ | 4.485 | 2.17  | 0.006    | 0.24   | 4.48    | 8.74    |
| ...     | ...   | ...   | ...      | ...    | ...     | ...     |
| $\gamma_{8,4}$ | 23.55 | 7.783 | 0.022    | 8.20   | 23.55   | 38.76   |
| $\gamma_{8,6}$ | 21.47 | 10.63 | 0.029    | 0.60   | 21.45   | 42.22   |

3.3.2. Normal Hierarchical Mixture Model.

The second alternative model involves the normal hierarchical of two mixtures with 200000 iterations, 100 thin, 200000 samples, 6 predictor variables at level 1 and 8 predictors at level 2. The estimation of this model includes 70 parameters in level 1 and 126 parameters in level 2.

The regression coefficients which were significant in the micro model were mostly found in mixture 2 of the hierarchical normal mixture. The total number of significant regression coefficients were 35 however Table 5 only illustrated the first 12 significant parameters. One of the parameters in Table 5, for example, $\beta_{1,9}$ explains the first parameter in the ninth province (Matabeleland South) in this study. This parameter indicates that it was significant as zero was excluded in the interval from 1.61 to 5.81.
Table 5 Significant parameters of Hierarchical Normal mixture in the micro model for mixture 1 and mixture 2

| No. | Parameter | mean | 2.5% | 97.5% | No. | Parameter | mean | 2.5% | 97.5% |
|-----|-----------|------|------|-------|-----|-----------|------|------|-------|
| 1   | $\beta_{0,1}$  | 72.2 | 63.09 | 81.39 | 7   | $\beta_{0,7}$  | 86.3 | 83.12 | 89.21 |
| 2   | $\beta_{0,2}$  | 78.6 | 67.09 | 82.2  | 8   | $\beta_{0,8}$  | 81.2 | 76.95 | 85.39 |
| 3   | $\beta_{0,3}$  | 96.5 | 86.77 | 108.2 | 9   | $\beta_{0,9}$  | 92.7 | 88.98 | 96.23 |
| 4   | $\beta_{0,4}$  | 91  | 88.49 | 93.83 | 10  | $\beta_{0,10}$ | 86  | 84.14 | 88.01 |
| 5   | $\beta_{0,5}$  | 86.6 | 85.43 | 87.87 | 11  | $\beta_{1,9}$  | 3.73 | 1.61 | 5.81  |
| 6   | $\beta_{0,6}$  | 85.8 | 83.92 | 87.59 | 12  | $\beta_{2,2}$  | -5.25 | -9.34 | -1.06 |

The predictors that significantly influence the pass rate of A-level national examination in mixture 1 includes the percentage of qualified teachers only. This variable was significantly witnessed in Harare and Mashonaland West provinces. The regression coefficients of the mixture 2 were all significant. This meant that the micro variables influencing the national examination are the public libraries, the teacher to pupil ratio and the classroom to teacher ratio. The number of public libraries significantly affects Matabeleland South and the number of orphaned and vulnerable learners significantly affects Harare, Mashonaland East, Mashonaland West, Masvingo, and Matabeleland South. Percentage of qualified teachers significantly affects Harare, Mashonaland East, Mashonaland West, Masvingo, and Matabeleland South. Teacher to pupil ratio significantly affects Bulawayo, Harare, Mashonaland Central, Mashonaland East, Mashonaland West, and Masvingo. The number of pupils on financial aid significantly affects Mashonaland East, Mashonaland West, and Matabeleland South. Lastly, classroom to teacher ratio significantly affects Bulawayo, Harare, Manicaland, Mashonaland East, and Midlands. Table 6 shows the significant regression coefficient parameters for mixture 1 and mixture 2 in the macro model.

Table 6. Significant parameters of Hierarchical Normal mixture in the macro model for mixture 2

| Parameter | mean | sd  | MC error | 2.50% | median | 97.50% |
|-----------|------|-----|----------|--------|--------|--------|
| $\gamma_{00}$ | 86.16 | 0.81 | 0.002 | 84.74 | 86.1 | 87.95 |
| $\gamma_{01}$ | 3.19  | 1.09 | 0.002 | 1.01  | 3.21  | 5.31  |
| $\gamma_{20}$ | 18.51 | 4.62 | 0.018 | 9.94  | 18.35 | 27.93 |
| $\gamma_{30}$ | 6.35  | 1.91 | 0.004 | 3.03  | 6.21  | 10.69 |
| $\gamma_{84}$ | 23.53 | 7.80 | 0.018 | 8.22  | 23.52 | 38.82 |
| $\gamma_{86}$ | 21.51 | 10.59 | 0.022 | 2.77  | 21.52 | 42.2  |

There were fifteen parameters that significantly affect the national examination under mixture 2 however only a few are shown in Table 6. In mixture 1, there were no parameters that significantly affect the national examination. All level 2 variables influence the national examination in mixture 2 and these include percentage of repeaters, the percentage of schools without electricity, the percentage of schools without water source, the average learners’ classroom ratio, the average learners per computer ratio, the number of pupils without seating facilities, the number of pupils without writing facilities and lastly school participation rate.

3.3.3. Log-logistic and Normal Hierarchical Regression Models.
The third and fourth alternative models include log-logistic and normal hierarchical regression models respectively. Six predictors in level 1 and eight predictors in level 2 were used in modelling of the A-level average pass rate and to identify the factors which influence the national examination. The
parameters which were estimated at the first level were 70 and at the second level, there were 63 parameters. For the log-logistic model, 182000 iterations and 100 thin were used to estimate the parameters with 182000 samples. On the other hand, for the normal model, the 226000 iterations with 100 thin and burn in of 5000 and 221001 samples were used to estimate the parameters.

Table 7 shows the regression coefficients found in log-logistic and normal models which significantly influence the national examination. Only five parameters were significant in the hierarchical log-logistic model and fifteen in the hierarchical normal model although only five are shown in Table 7. Under the hierarchical log-logistic model, the only variables that influence the A-level national examination are public libraries and were observed in Mashonaland East. On the other hand, under the hierarchical normal model, the pass rate of A-level national examination is significantly affected by the public libraries, the teacher to pupil ratio, the pupils who depend on financial aid and the classroom to teacher ratio. The micro model regression coefficients which were significant were only found in Mashonaland East and Bulawayo provinces. Mashonaland East province is significantly affected by the number of teachers to pupil ratio, pupils who depend upon financial aid and classroom to teacher ratio and public libraries influence the pass rate in Bulawayo province. As for the macro model, the significant regression coefficients are shown in Table 8.

| Parameter  | Mean | SD  | MC Error | 2.5%  | Median | 97.5% |
|------------|------|-----|----------|-------|--------|-------|
| $\beta_{0.4}$ | 2.26 | 0.68 | 0.003    | 0.52  | 2.38   | 3.28  |
| $\beta_{0.5}$ | 2.00 | 0.62 | 0.006    | 0.67  | 2.04   | 3.13  |
| $\beta_{0.6}$ | 2.54 | 0.95 | 0.005    | 0.30  | 2.73   | 3.96  |
| $\beta_{0.9}$ | 2.22 | 0.98 | 0.005    | 0.07  | 2.31   | 3.86  |
| $\beta_{0.10}$ | 1.28 | 0.62 | 0.003    | 0.03  | 1.29   | 2.48  |

| Parameter  | Mean | SD  | MC Error | 2.5%  | Median | 97.5% |
|------------|------|-----|----------|-------|--------|-------|
| $\gamma_{00}$ | 83.4  | 2.71 | 0.03     | 80.63 | 83.53  | 85.97 |
| $\gamma_{05}$ | 17.53 | 8.93 | 0.02     | 0.05  | 17.55  | 34.93 |
| $\gamma_{40}$ | 5.41  | 2.57 | 0.02     | 0.48  | 5.47   | 10.05 |
| $\gamma_{75}$ | -237.2 | 93.25 | 0.20     | -420.3 | -237.1 | -54.75 |
| $\gamma_{85}$ | -53.11 | 20.51 | 0.05     | -93.22 | -53.17 | -13.17 |

Table 8. Significant parameters of the hierarchical macro models
(a) Log-logistic

| Parameter  | Mean  | SD  | MC Error | 2.5%  | Median | 97.5% |
|------------|-------|-----|----------|-------|--------|-------|
| $\gamma_{05}$ | 17.52 | 8.90 | 0.02     | 0.06  | 17.52  | 34.97 |
| $\gamma_{85}$ | -53.15 | 20.53 | 0.04     | -93.23 | -53.16 | -12.73 |
The macro model significant parameters for log-logistic were only five whilst for normal they were two. There were only three variables that positively influenced the national examination these include the percentage of schools without a water source, the number of pupils without writing facilities and the school participation rate.

4. Selection of the Best Model
Table 9 shows the DIC values used in this study for model fitting. The smallest value of the DIC indicated the best model to fit the A-Level national examination pass rate in Zimbabwe and in this study, it is the hierarchical normal mixture model.

| Model                                      | DIC  |
|--------------------------------------------|------|
| Alternative 1 - Hierarchical Logistics Mixture | 913.88 |
| Alternative 2 - Hierarchical Logistics     | 419.46 |
| Alternative 3 - Hierarchical Normal Mixture | 389.34 *** |
| Alternative 4 - Hierarchical Normal        | 428.31 |

5. Modelling of the A-level national examination

The best model influencing the national examination was modelled and its micro models is shown by equation (15). For example, the micro model for Bulawayo province in the hierarchical normal mixture can be written as follows:

\[
y_{1,1} = 0.08(-83.3 + 538.7X_{1,1,1} - 1.01X_{2,1,1} - 1.01X_{3,1,1} - 83.3X_{4,1,1} + 538.8X_{5,1,1} - 1.01X_{6,1,1}) + 0.92(72.19 - 0.04X_{1,2,1} + 1.27X_{2,2,1} + 0.96X_{3,2,1} - 5.2X_{4,2,1} - 0.18X_{5,2,1} + 2.61X_{6,2,1}) \tag{15}
\]

The micro model for each province can be written in the same way as in equations (15). In modelling the combined hierarchical normal mixture for the average pass rate of the A-Level national examination in Zimbabwe, the best model was modelled as shown in equation (16) where \( \mu_p \) is the combined mixture model.

\[
\mu_p = 0.08(7.63 + 239.8X_1 + 0.56X_1 + 7.34G_4 + 10.66G_8 + 60.13G_1X_1 + 372.8G_8X_1 + 2.83G_1X_6 + 1.39G_8X_6) + 0.92(86.16 - 0.36X_1 + 6.36X_6 + 1.05G_1 + 5.38G_8 + 3.19G_1X_1 + 1.59G_8X_2 - 0.55G_1X_2 + 1.45G_8X_2 + 2.07G_1X_6 + 21.51G_8X_6) \tag{16}
\]

6. Conclusion
In summary, it can be concluded that the 2017 A-level average pass rate can be modelled using a hierarchical normal mixture model. There were variations of regression coefficients of the micro and macro model between provinces. This variation proved to be significantly influenced by district and province variables. The factors which influence the national examination percentage pass rate under the micro model in mixture 1 was only the percentage of qualified teachers whilst in mixture 2, the variables include the public libraries, the teacher to pupil ratio and the classroom to teacher ratio. At macro level in mixture 1, all level 2 predictors were not significant however in mixture 2, all level 2 characteristics influence the national examination and these include the percentage of repeaters, the percentage of schools without electricity, the percentage of schools without water source, the average learners classroom ratio, the average learners per computer ratio, the number of pupils without seating facilities, the number of pupils without writing facilities and lastly the school participation rate.
References

[1] Kellaghan T, and Greaney V 2004 Monitoring performance: Assessment and examinations in Africa. Paris: Association for the Development of Education in Africa; Washington DC: World Bank

[2] Hox J J, Moerbeek M, and Van de Schoot R 2010 Multilevel Analysis: Techniques and Applications, Quantitative Methodology Series. New (York: Routledge)

[3] Iriawan N 2012 Modelling and Data-Driven Analysis, Volume 1 (Surabaya: ITS Press)

[4] McLachlan G, and Basford K 1988 Mixture Models: Inference and application to Clustering. Marcel and Decker Inc

[5] Guure C, Ibrahim N A, Dwamah, D, and Bosomprah S 2014 Bayesian statistical inference of the log-logistic model with interval-censored lifetime data. Journal of Statistical Computation and Simulation, 4.

[6] Ismartini P, Iriawan N, and Ulama B S S 2013 Comparison of Unilevel and Multilevel Model for Hierarchical Structure Data Analysis Using Bayesian Approach. Paper Presented at The Proceeding of the Fourth National Mathematics Conference

[7] Raudenbush S W and Bryk A S 2002 Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition (Vol. 1). (London: Sage)

[8] Goldstein H 1995 Multilevel Statistical Models (London: Edward Arnold)

[9] Zulvia P 2017 Pemodelan multilevel dan analisis data panel pada penelitian Pendidikan: Studi Kasus- Data Ujian Nasional SMA di Jawa Barat (Thesis), Bogor

[10] Carlin B P and Chib, S 1995 Bayesian Model Choice via Markov Chain Monte Carlo Methods. Journal of the Royal Statistical Society. Series B (Methodological), 473-484

[11] Ntzoufras I 2009 Bayesian Modeling in WinBugs (New Jersey, USA: John Wiley and Sons, Inc)

[12] Box G E and Tiao G C 1973 Bayesian Inference in Statistical Analysis (Massachusetts: Addison Wesley)

[13] Richardson S, and Green P J 1997 On Bayesian Analysis With an Unknown Number of Components, Journal of Royal Statistical Society, 59(4), 731-792

[14] Iriawan N 2000 Computationally intensive approaches to Inferences in Neo-Normal Linear Models (Thesis), CUT-Australia.

[15] Susanto, I., Iriawan, N., Kuswanto, H., Suhartono, Fithriasari, K., Ulama, B.S.S., Suryaningtyas, W., and Pravitasari, A.A. 2018 On the Markov Chain Monte Carlo Convergence Diagnostic of Bayesian Finite Mixture Model for Income Distribution, Journal of Physics: Conference Series, 1090(1), 012014, DOI: 10.1088/1742-6596/1090/1/012014.

[16] Koop G 2003 Bayesian Econometrics. (Manchester: John Wiley and Sons)

[17] Chen C, Zhang G, Liu X C, Ci Y, Huang H, and Ma J 2016 Driver injury severity outcome analysis in rural interstate highway crashes a two-level Bayesian logistics regression interpretation. Elsevier, 73

[18] Spiegelhalter D J, Best N G, Carlin B P, and Van der Linder A 2002 Bayesian Measures of model complexity and fit (with discussions). Journal of Royal Statistics, 62,583-639